



UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

La Universidad Católica de Loja

ÁREA TÉCNICA

TÍTULO DE INGENIERO EN INFORMÁTICA

**Análisis de los usos de la tecnología en los estudiantes de la
Universidad Metropolitana**

TRABAJO DE TITULACIÓN

AUTOR: Jiménez Gaona, Juan Paúl

DIRECTORA: Jara Roa, Dunia Inés, Mgs.

CENTRO UNIVERSITARIO CARIAMANGA

2016

APROBACIÓN DE LA DIRECTORA DEL TRABAJO DE TITULACIÓN

Magister

Dunia Inés Jara Roa

DIRECTORA DEL TRABAJO DE TITULACIÓN

De mi consideración:

El presente trabajo de titulación: **Análisis de los usos de la tecnología en los estudiantes de la Universidad Metropolitana** realizado por **Jiménez Gaona Juan Paúl**, ha sido orientado y revisado durante su ejecución, por cuanto se aprueba la presentación del mismo.

Loja, mayo 2016.

f)

DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS

“Yo, **Jiménez Gaona Juan Paúl** declaro ser autor del presente trabajo de titulación: **Análisis de los usos de la tecnología en los estudiantes de la Universidad Metropolitana**, de la Titulación de **Ingeniero en Informática**, siendo **Dunia Inés Jara Roa** directora del presente trabajo; y eximo expresamente a la Universidad Técnica Particular de Loja y a sus representantes legales de posibles reclamos o acciones legales. Además certifico que las ideas, conceptos, procedimientos y resultados vertidos en el presente trabajo investigativo, son de mi exclusiva responsabilidad.

Adicionalmente declaro conocer y aceptar la disposición del Art. 88 del Estatuto Orgánico de la Universidad Técnica Particular de Loja que en su parte pertinente textualmente dice: “Forman parte del patrimonio de la Universidad la propiedad intelectual de investigaciones, trabajos científicos o técnicos y tesis de grado que se realicen a través, o con el apoyo financiero, académico o institucional (operativo) de la Universidad”

f.

Jiménez Gaona Juan Paúl

110469548-9

DEDICATORIA

A DIOS, que gracias a él tengo la fortaleza y salud para seguir adelante frente a cualquier adversidad.

Para mi Madrecita Sra. Carmen Zenaida Gaona Rojas, quien en todo momento me ha brindado su amor incondicional, apoyo moral y económico para lograr este sueño tan anhelado, a ti querida madrecita.

A mis hermanos: en especial a Cosme, Carlos, Claudio y Cecilia, quienes siempre han estado presente brindándome su apoyo emocional y económico para cristalizar mis sueños, mil gracias.

A mis queridos sobrinos: Johanna, Jonathan, Jelissa, Daniel, Cristhoper y a nuestra pequeña Dayanna, el nuevo regalo de Dios en nuestra familia.

A mis queridos abuelitos: María y Emiliano; a mi inolvidable y querida bisabuelita: María (+) y a todos mis familiares.

Paúl

AGRADECIMIENTO

En primer lugar a Dios y la Virgen del Cisne, por brindarme su cobijo espiritual, salud y fortaleza.

A mis querida familia: por todos su apoyo emocional y económico, mil gracias: mamita, Cosme, Carlos, Claudio, Cecilia, Fanny y Lady, por alentarme a continuar en mi vida profesional y guiarme por el sendero correcto y estar presentes en cada instante de mi vida con su apoyo único e incondicional.

A mi querida institución que me formo y me abrió las puertas al mundo del saber la Universidad Técnica Particular de Loja y a cada uno de mis docentes y personal de apoyo. Y de manera muy especial a la Mgs. Dunia Inés Jara Roa, por sus conocimientos compartidos, su paciencia y apoyo para culminar con el presente trabajo.

De igual manera para el Dr. Juan Carlos Torres, Director General del presente proyecto, el cual con sus conocimientos impartidos y sabios consejos me ha permitido llegar a esta instancia.

Así mismo, a la Lcda. Lidia Villacis, quién siempre estuvo presente en cada momento de mi formación profesional y brindándome el apoyo y guía para cumplir con mis propósitos académicos.

A mis queridas amigas y amigos, mis apreciados y entrañables compañeros de aula con quienes tuve el gusto de compartir penas y alegrías durante nuestra formación académica.

A todos, mil gracias por cada una de esas palabras de aliento y apoyo incondicional que me sirvieron para seguir adelante y no decaer en este gran reto.

Mis sinceros agradecimientos a todos y cada uno de ustedes, gracias infinitas.

Paúl

ÍNDICE DE CONTENIDOS

CARATULA	I
APROBACIÓN DE LA DIRECTORA DEL TRABAJO DE TITULACIÓN	II
DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS.....	III
DEDICATORIA	IV
AGRADECIMIENTO	V
ÍNDICE DE CONTENIDOS	VI
ÍNDICE DE TABLAS	X
ÍNDICE DE FIGURAS.....	XI
RESUMEN.....	1
ABSTRACT	2
INTRODUCCIÓN.....	3
1.1. Introducción.....	4
1.2. Objetivos	5
1.3. Preguntas de investigación.....	5
1.4. Hipótesis.....	6
MARCO TEÓRICO	7
2.1. Brecha digital.....	8
2.1.1. Determinantes de la brecha digital.....	8
2.1.2. La brecha digital en Latinoamérica	10
2.1.3. Brecha digital en el Ecuador	11
2.2. Usos de Internet	12
2.2.1. Académico.....	12
2.2.2. Comunicación.....	15
2.2.3. Entretenimiento	16
2.3. Rendimiento académico	16
2.4. Minería de datos.....	18
2.4.1. Tipos de datos	18

2.4.2.	Tipos de modelos	19
2.4.3.	Fases del proceso de minería de datos	19
METODOLOGÍA		22
3.1.	Población y muestra	23
3.2.	Integración y recopilación	24
3.3.	Selección, limpieza y transformación	26
3.4.	Minería de datos	27
3.5.	Tarea de minería de datos a utilizar: clusterización	28
3.6.	Análisis discriminante	29
3.7.	Comprobación de hipótesis: regresión logística	29
RESULTADOS		31
4.1.	Descripción de la muestra	32
4.2.	Acceso a Internet: lugar de conexión, tiempo y nivel de conocimiento	34
4.3.	Usos de Internet en el ámbito académico y de entretenimiento	37
4.4.	Uso de redes sociales	46
4.5.	Incidencias entre variables	46
4.6.	Factorización: reducción de variables	48
4.7.	Clusterización: determinación de perfiles	48
4.7.1.	Perfil académico	48
4.7.2.	Perfil de entretenimiento	50
4.8.	Descripción de perfiles	50
4.8.1.	Académico	50
4.8.2.	Entretenimiento	52
4.9.	Rendimiento académico	53
4.10.	Comprobación de hipótesis	54
4.10.1.	El nivel de ingresos determina como se utiliza Internet para el aprendizaje	54
4.10.2.	El nivel de ingresos determina como se utiliza Internet para entretenimiento	55
4.10.3.	El uso de la tecnología en el aprendizaje incide en el rendimiento académico	56

- 4.10.4. El uso de la tecnología para entretenimiento incide en el rendimiento académico.
56

ANÁLISIS Y DISCUSIÓN DE RESULTADOS.....	58
5.1. Análisis hipótesis 1	59
5.2. Análisis hipótesis 2	59
5.3. Análisis hipótesis 3	60
5.4. Análisis hipótesis 4	61
CONCLUSIONES Y RECOMENDACIONES	63
Conclusiones	64
Recomendaciones	65
BIBLIOGRAFÍA.....	66
ANEXOS.....	69
Anexo 1	70
Anexo 2	72
Anexo 3	73
Anexo 4	74
Anexo 5	75
Anexo 6	76
Anexo 7	77
Anexo 8	78
Anexo 9	79
Anexo 10	80
Anexo 11	81
Anexo 12	82
Anexo 13	83
Anexo 14	84
Anexo 15	85
Anexo 16	86
Anexo 17	87

Anexo 18	88
Anexo 19	89
Anexo 20	90
Anexo 21	91
Anexo 23	93
Anexo 24	94
Anexo 25	95
Anexo 26	96
Anexo 27	97
Anexo 28	98
Anexo 29	100
Anexo 30	102
Anexo 31	104
Anexo 32:	106
Anexo 33	112

ÍNDICE DE TABLAS

Tabla 1: Transformación de variables	26
Tabla 2: Generación de nueva variable rendimiento académico	26
Tabla 3: Determinación de variables para comprobación de hipótesis	27
Tabla 4: Ingresos semanales a la plataforma virtual	37
Tabla 5: Consultas mensuales al profesor	38
Tabla 6: Consultas mensuales a compañeros	39
Tabla 7: Medianas de las variables académicas	49
Tabla 8: Exactitud de clasificación de grupos académicos.....	49
Tabla 9: Porcentaje de exactitud de grupos de uso de Internet para entretenimiento	50
Tabla 10: Grupos de rendimiento académico.....	53
Tabla 11: Coeficiente de regresión logística para los ingresos y el uso de Internet en actividades académicas.....	54
Tabla 12: Coeficiente del modelo de regresión logística para los ingresos y los usos de Internet en entretenimiento	55
Tabla 13: Coeficiente de regresión logística entre usos de Internet para actividades académicas y rendimiento académico	56
Tabla 14: Coeficiente de regresión logística para usos de Internet en entretenimiento y rendimiento académico.....	57

ÍNDICE DE FIGURAS

Figura 1: Determinantes del rendimiento académico	17
Figura 2: Distribución de los alumnos por género	32
Figura 3: Distribución de alumnos por carrera.....	33
Figura 4: Distribución por edades	33
Figura 5: Ingresos mensuales de las familias de los alumnos.....	34
Figura 6: Distribución de los alumnos según el lugar de conexión	34
Figura 7: Días de conexión a la semana.....	35
Figura 8: Nivel de conocimiento en el manejo de Internet.....	35
Figura 9: Distribución de los alumnos por horas de conexión	36
Figura 10: Años de experiencia en el uso de Internet	36
Figura 11: Descarga de recursos educativos de la plataforma virtual	40
Figura 12: Videos académicos visualizados en YouTube mensualmente	40
Figura 13: Foros virtuales mensuales	41
Figura 14: Post o Tweets académicos en redes sociales mensuales.....	41
Figura 15: Horas de chat académico mensuales	42
Figura 16: Horas de búsqueda de información académica en Internet por mes.....	42
Figura 17: Horas de acceso a la biblioteca virtual de la universidad	43
Figura 18: Horas de chat por diversión a la semana	44
Figura 19: Horas semanales de uso de redes sociales.....	44
Figura 20: Horas semanales para descargar música, videos y programas	45
Figura 21: Videos de YouTube para diversión semanalmente	45
Figura 22: Seguidores, amigos o contactos en redes sociales.....	46
Figura 23: Centroides de los grupos del uso de Internet en actividades académicas.....	51
Figura 24: Centroides de los grupos del uso de Internet para entretenimiento.....	52

RESUMEN

El presente trabajo de titulación, tiene como objeto el estudio de los ingresos económicos de los alumnos de la Universidad Metropolitana de Quito y la incidencia en los usos de Internet para actividades académicas y de entretenimiento; de la misma manera el nivel de incidencia de los usos de Internet en actividades académicas y de entrenamiento sobre el rendimiento académico de los estudiantes.

Para lograr el objetivo del proyecto de titulación y comprobar las hipótesis planteadas se hizo uso de técnicas de minería de dato como: análisis clúster y discriminante; y regresión logística como modelo de minería de datos.

PALABRAS CLAVE: brecha digital, ingresos económicos, usos de Internet, recopilación, limpieza de datos, clusterización, regresión, minería de datos, rendimiento académico, académico, entretenimiento

ABSTRACT

This work degree, aims to study the income of students of the Metropolitan University of Quito and the impact on Internet applications for academic activities and entertainment; in the same way the level of incidence of Internet use in academic activities and training on the academic performance of students.

To achieve the project objective titling and check the hypotheses made use of data mining techniques such as cluster and discriminant analysis; and logistic regression as data mining model.

KEYWORDS: digital divide, income, Internet applications, data collection, data cleansing, clustering, regression, data mining, academic performance, academic, entertainment

CAPÍTULO I
INTRODUCCIÓN

1.1. Introducción

En la actualidad el avance de la sociedad circula en torno del desarrollo de nuevas herramientas y dispositivos tecnológicos de comunicación. El Internet, se ha convertido en el principal aliado para el progreso de muchas naciones en todos sus ámbitos de desarrollo.

La presente investigación, parte del estado actual de los usos de Internet en las Universidades del Ecuador para actividades académicas y de entretenimiento por parte de los alumnos, y como esto influye en el nivel académico de los mismos. Además, se determina la brecha digital existente entre los estudiantes de diferentes niveles económicos. También se enmarca en determinar cuál es la incidencia de los ingresos económicos familiares sobre el uso de Internet tanto en actividades académicas como de entretenimiento, así mismo de qué forma influye el uso de Internet en actividades académicas y de entretenimiento y su nivel de incidencia en el rendimiento académico de los alumnos de la Universidad Metropolitana del Ecuador (Quito).

Haciendo uso de técnicas de minería de datos se recolectó, seleccionó, limpió, relacionó y se logró comprobar la validez de las hipótesis del trabajo de investigación.

A continuación se describe los cinco capítulos que conforman el trabajo de investigación:

Capítulo I: está constituido por los objetivos del proyecto, las preguntas de la investigación, las hipótesis a comprobar.

Capítulo II: se realiza un estudio de la brecha digital, los usos de Internet y la extracción de conocimiento haciendo uso de técnicas de minería de datos.

Capítulo III: se detalla la metodología seguida como son: cálculos de población y muestra; técnicas de recolección de información: encuestas, entrevistas y las técnicas de minerías de datos: factorización, clusterización, análisis discriminantes y regresión logística.

Capítulo IV: se visualiza los resultados obtenidos siguiendo la metodología descrita en el capítulo III, como son: resultados de las encuestas, relación existente entre los niveles de ingresos y los perfiles de los alumnos; así como, los perfiles de alumnos con el rendimiento académico.

Capítulo V: se presenta el análisis y discusión de la información descubierta.

Por último se concluyó y recomendó en base a las observaciones del presente trabajo.

1.2. Objetivos

- Determinar los usos de Internet en actividades académicas y de entretenimiento de la Universidad Metropolitana.
- Determinar del rendimiento académico de los estudiantes de la Universidad Metropolitana.
- Relacionar los niveles ingresos con el uso de Internet para actividades académicas y rendimiento académico.
- Relacionar el uso de Internet en actividades académicas y el rendimiento académico.
- Relacionar el uso de Internet para entretenimiento y el rendimiento académico.

1.3. Preguntas de investigación

- ¿Cómo se relacionan los niveles de ingreso de las familias de los estudiantes universitarios con los usos de Internet en actividades académicas y de entretenimiento?
- ¿Cómo se relacionan el rendimiento académico y los usos de Internet en actividades académicas y de entretenimiento?

1.4. Hipótesis

Hipótesis relacionadas con la interrogante N° 1:

Hipótesis 1: el nivel de ingresos determina como se utiliza Internet para el aprendizaje.

Hipótesis 2: el nivel de ingresos determina como se utiliza Internet para entretenimiento.

Hipótesis relacionadas con la interrogante N° 2:

Hipótesis 3: el uso de la tecnología en el aprendizaje incide en el rendimiento académico.

Hipótesis 4: el uso de la tecnología para entretenimiento incide en el rendimiento académico.

CAPÍTULO II
MARCO TEÓRICO

En el presente capítulo se define la brecha digital y los determinantes que influyen en ella como: económico, demográfico y cultural. También la brecha digital en Latinoamérica y en Ecuador; así como el uso de Internet en sus ámbitos académicos, de comunicación, entretenimiento y rendimiento académico.

2.1. Brecha digital

“La premisa básica de la brecha digital es la diferencia que existe entre individuos y sociedades que tienen acceso a los recursos tecnológicos de cómputo, telecomunicaciones e Internet” (Gallardo, 2006).

Serrano y Martínez (2003), definen a la brecha digital como la disociación que existe entre personas, comunidades, estados, etc., que hacen uso de Tecnologías de la Información y Comunicación (TIC) como parte fundamental de su diario vivir y aquellas que no tienen acceso a las mismas y aunque tengan acceso no tienen el conocimiento básico para utilizarlas.

De tal manera que la brecha digital se puede definir como la desigualdad de acceso a las TIC de un individuo o sociedad, esta desigualdad puede estar definida por varios factores, como: sociales, políticos, culturales o económicos. Además la brecha digital se ve determinada en qué forma se usa la información y la utilización adecuada de la misma.

2.1.1. Determinantes de la brecha digital

La brecha digital se acrecienta o disminuye según sus determinantes que inciden de una u otra manera en la desigualdad digital de un individuo o una sociedad en particular. Entre los determinantes que más inciden en la brecha digital se puede describir los siguientes:

- Determinante económico

“La noción de “brecha digital” refleja una profunda desigualdad entre las naciones ricas y las pobres en el acceso a Internet como un recurso económico, político y cultural de vastas consecuencias” (Sar, 2004).

Es así que el aspecto económico no solo incide en la determinación de mayor acceso al sector rico de una sociedad, sino también delimita las oportunidades de los sectores menos

favorecidos económicamente para acceder a herramientas tecnológicas y a la información digital.

Gallardo (2006), expresa que muchos estados que no poseen avances técnicos, de infraestructura y capacidad institucional, se han visto en la obligación de estar rezagados del resto de países del mundo, y el problema surge no solo en base de que la tecnología beneficia a los países ricos, sino que también impone más límites de desarrollo a los países pobres.

Los niveles de ingresos económicos de un individuo o sociedad siempre van a implicar directamente en acercar o distanciar la desigualdad digital, como se puede ver en varios estudios realizados la cantidad de ingresos puede determinar qué cantidad de recursos se distribuyen para adquirir y acceder a herramientas tecnológicas.

- **Determinante demográfico**

El determinante demográfico de la brecha digital se lo puede determinar en base al género, edad y sociedad del individuo. Gallardo (2006), manifiesta que en algunas culturas y zonas geográficas del mundo, la brecha digital determinada por el género es muy estrecha; en cambio en otras zonas donde la mujer es objeto de discriminación o marginación las brechas digitales, sociales, económicas y culturales tienden a ser más amplias con relación al género masculino.

El grado de determinación del género en la brecha digital se ve enmarcado según factores sociales y culturales, en nuestra sociedad se puede observar la igualdad de género en acceso a información digital y oportunidades laborales referentes a tecnología.

La medida en que el capital humano es fomentado, empleado, y reconocido es profundamente social y con frecuencia se ha examinado a lo largo de las líneas de género. Existe amplia literatura que examina cómo el desarrollo y el reconocimiento de habilidades cuantitativas se traduce en desigualdades educativas y ocupacionales para hombres y mujeres (Hargittai y Shafer, 2006).

La brecha digital puede también estar determinada por el nivel académico de un individuo o sociedad. Además las destrezas y habilidades innatas o desarrolladas de una persona, pueden determinar la manera en como accede a la información y el aprovechamiento de la misma (Torres, 2012).

- **Determinante cultural**

Entre los factores que inciden en la determinación cultural de la brecha digital se tiene el idioma.

La cuestión del uso de los diferentes idiomas utilizados en Internet es un asunto que aunque parece exclusivamente lingüístico, no lo es. El idioma es uno de los más complejos y significativos temas relacionados con el contenido y con otros argumentos más amplios de la tecnología de la información. El idioma se relaciona con muchas formas de división social que a su vez se vinculan con la nacionalidad, la economía, la cultura, la educación y la lectura; afecta dramáticamente la posibilidad de diversos grupos de tener acceso a Internet y publicar en ella en la medida en que es un medio de expresión de las diferentes identidades culturales (Gallardo, 2006).

Con lo anteriormente citado se puede observar que el idioma incide drásticamente en como una persona accede a la información digital y hacer uso de las TIC adecuadamente. Como es de conocimiento general, la mayoría de información publicada digitalmente se encuentra en idioma inglés, el cual hoy en día es el idioma universal de información y negocios. Por lo tanto, el idioma determina como la información digital es publicada y aprovechada.

2.1.2. La brecha digital en Latinoamérica

Como se conoce en Latinoamérica no todos los países tienen la misma infraestructura tecnológica, ni brinda las facilidades para que sus ciudadanos accedan con facilidad y aprovechen las TIC. En muchos países de Latinoamérica las políticas no favorecen a las personas para que accedan a Internet de forma fácil y así estar acorde al avance tecnológico que hoy en día se ha convertido en una globalización tecnológica.

Según el reporte de Global Information Technology, Weforum (2014), en Latinoamérica el país que más avanza en lo referente a implementación de infraestructura tecnológica y con mayor acceso y utilización de las TIC es Chile, cabe recalcar que a pesar del avance que han tenido los países de América Latina en materia tecnológica, aún existe una gran brecha comparada con regiones como Europa, Asia y Estados Unidos, en donde la infraestructura es más avanzada y el uso de la TIC es aprovechado de mayor y mejor manera.

2.1.3. Brecha digital en el Ecuador

En el Ecuador en los últimos años se ha incrementado drásticamente la infraestructura tecnológica para poder acceder y aprovechar las TIC. En la publicación más reciente de Weforum (2014), Ecuador logra el puesto veintinueve (29) de 148 países analizados en lo referente a adquisición de tecnología avanzada; en lo concerniente a acceso a Internet en las escuelas ocupa el puesto ochenta (80); y en la calidad del sistema educativo ocupa el puesto sesenta y dos (62).

Con los datos referenciados se puede determinar que el Ecuador ha sufrido cambios en aspectos como: nivel de educación que brinda, la infraestructura tecnológica y el desarrollo del pensamiento de sus ciudadanos para involucrarlos en la sociedad de la información digital.

Este avance en materia tecnológica, en los últimos años en el Ecuador se lo puede confirmar con datos expuestos por el INEC (Instituto Nacional de Estadísticas y Censos) en el último informe del Inec (2013), se observa, que en el 18,1% de los hogares ecuatorianos tienen mínimo un computador portátil, lo cual representa un incremento del 9,1% más con referencia al año 2010 (en donde solo el 9% de los hogares contaban con este dispositivo tecnológico en casa).

En el mismo estudio muestra que el 27,5% de los hogares del país cuentan con un computador de escritorio, en donde se ve reflejado un incremento del 3,5% con referencia al año 2010, en donde solo el 24% de los hogares contaban con un computador de escritorio.

En lo referente al acceso a Internet, se puede observar que, el 28,3% de los hogares ecuatorianos tienen acceso a Internet, es decir un 16,5% más con referencia al año 2010 en donde el porcentaje de hogares que tenían acceso a Internet era del 11,8%. El porcentaje de crecimiento del acceso a Internet en el sector urbano es de 20,3% y en el sector rural es de 7,8%, es decir el sector urbano crece 12,5% más rápido que el rural, tanto en lo que es infraestructura y servicios tecnológicos.

Así mismo, se determina que del 100% de hogares que tienen acceso a Internet, el 32% utiliza Internet como fuente de información; el 31,7% como medio de educación y aprendizaje; el 25,5% como comunicación en general; el 4,9% por cuestiones de trabajo y el 5,9% restante aprovechan Internet para otros ámbitos.

2.2. Usos de Internet

El uso de Internet se ha convertido en una costumbre indispensable, en la actualidad el Internet es otro servicio básico que se debe disponer, el mismo que se lo puede utilizar como medio de: información, educación, comunicación, trabajo o simplemente como un medio de entretenimiento.

Se dice que la Internet ampliará el acceso a la educación, buenos empleos y una mejor salud; y que va a crear nuevos espacios de deliberación para la discusión política y ofrecer a los ciudadanos el acceso directo al gobierno (DiMaggio y Hargittai, 2001).

De tal manera que el Internet ha abierto una gran gama de beneficios en todos los ámbitos que ha sido empleado. Cómo se puede observar Internet brinda facilidad de acceso a información digital de forma rápida y eficiente, y brindando ahorro de tiempo, esfuerzo y dinero.

Cabe recalcar que, el uso del Internet no solo se basa en si tiene acceso o no al mismo, sino las capacidades y destrezas que cada individuo tiene para acceder y hacer uso de las TIC de forma adecuada.

El uso del Internet puede ir orientado al ámbito académico, a las comunicaciones y al entretenimiento, como se describe a continuación:

2.2.1. Académico

Internet se ha convertido en un aliado en el proceso de enseñanza-aprendizaje, ya que gracias a él se puede acceder y compartir información. Dentro del uso del Internet en el ámbito académico se han desarrollado varias herramientas para poder acceder a la información de forma rápida, veraz y confiable, como: bibliotecas virtuales, revistas digitales, google scholar, etc. Para poder aprovechar estas herramientas es fundamental contar con destrezas digitales.

Para utilizar los recursos de Internet, debe tenerse como premisa el fomento en los usuarios de una actitud crítica hacia la información circulante. El volumen y la diversa confiabilidad de los productores de la información que se hallan en Internet ponen en primer plano la cuestión de la capacidad del usuario para diferenciar la información relevante y veraz de la que no lo es. Sólo la construcción de criterios de selección, el desarrollo de destrezas de alto orden y la adquisición de valores, permitirá a los usuarios seleccionar la información. Sin

duda, la falta de conocimiento sobre una temática, complica la búsqueda de nueva información hasta transformarla en un laberinto de difícil salida (Ilabaca, s.f.).

No es suficiente con tener acceso a la información digital, sino también la forma en como esta es discernida en base a un tema de interés. La información compartida en Internet no siempre es confiable. Por lo tanto, aquí se emplea las destrezas digitales del internauta para: buscar, discrepar, analizar y aplicar la información digital de manera adecuada. Es así, que se genera la calidad de información dependiendo del individuo que hace uso de ella.

La educación virtual representa un fenómeno de estudio relevante para comprender los desafíos que actualmente tiene la sociedad frente a los procesos que han generado los avances científicos y los desarrollos tecnológicos. En el campo educativo es indispensable estudiar las repercusiones que produce Internet en la educación virtual, las propuestas teóricas que se discuten actualmente respecto a interactividad y navegación, el desarrollo de objetos de aprendizaje, y las perspectivas sociales e institucionales que se esperan en este siglo que inicia de la educación basada en tecnologías de red (Garduño, 2007).

Los objetivos de la educación ya no solo se enmarcan a la impartición de conocimientos dentro de un salón de clases, sino también enmarcada a la educación en entornos virtuales, en donde el docente puede administrar los conocimientos a compartir con sus alumnos, haciendo uso de herramientas tecnológicas y los alumnos pueden acceder a una infinidad de información de interés para su educación académica.

E-learning

En lo referente al uso de Internet en el ámbito académico, se tiene las plataformas e-learning como: los Sistemas de Manejo de Contenidos (CMS) y los entornos virtuales de aprendizaje o Learning Management System (LMS).

Los CMS en los últimos años han progresado en tres etapas evolutivas, que han impactado, cada vez de forma más notoria, sobre la velocidad de creación de contenidos, el coste, la flexibilidad, la personalización del aprendizaje, la calidad en la atención del estudiante y las ventajas competitivas de las organizaciones que han aplicado las soluciones de e-learning (Boneu, 2007).

La evolución de las e-learning empieza con los CMS, permite crear páginas dinámicas y la gestión de contenidos en línea, pero tienen la limitante que no poseen herramientas colaborativas ni ayuda en tiempo real.

En base a los CMS surgen los LMS, los cuales ya permite la colaboración de multiusuario para: mantenimiento, actualización y ampliación de la web. En este tipo de herramientas ya surgen las herramientas de manejo de contenidos a nivel académico, y permite personalizar las herramientas a nivel institucional o del usuario.

Por último los LCMS (Sistemas de Gestión de Contenidos de Aprendizaje), estas herramientas están integradas por los CMS y los LCMS, de las cuáles hereda sus funcionalidades, permitiendo de forma eficiente la creación y administración de contenidos.

LMS, es un software instalado generalmente en un servidor web (puede instalarse en una intranet), que se emplea para crear, aprobar, administrar, almacenar, distribuir y gestionar las actividades de formación virtual (puede utilizarse como complemento de clases presenciales o para el aprendizaje a distancia). Un LMS se centra en gestionar contenidos creados por una gran variedad de fuentes diferentes. La labor de crear los contenidos para los cursos es desarrollada mediante un LCMS (Castro, Clarenc, Lenz, Moreno y Tosco, 2013).

La relación entre usuario de un LMS puede ser docente \leftrightarrow alumno; alumno \leftrightarrow alumno; alumno \leftrightarrow docente; docente-alumno \rightarrow contenido. Es así, que los contenidos son utilizados para estructurar los cursos, el docente hace uso de esos contenidos para afianzar su sistema de enseñanza-aprendizaje y por último el alumno accede a estos contenidos para ampliar sus conocimientos, desarrollando talleres o actividades de acuerdo a la temática del curso.

Los LMS no solo permiten administrar el contenido en un proceso de enseñanza-aprendizaje virtual, sino también el avance de los alumnos y evaluar las destrezas y conocimientos adquiridos. Los contenidos de un curso virtual son creados con un LCMS.

No se puede dejar a un lado que el Internet se ha vuelto un aliado indispensable para el estudio superior tanto presencial como a distancia, ya que gracias a este recurso se puede autoeducarse y formar parte de herramientas colaborativas en donde el tutor y estudiantes pueden interactuar de tal manera que acceden a nuevos recursos en línea.

Los entorno virtuales de aprendizaje, permiten tener acceso a herramientas colaborativas de conocimientos para los procesos de enseñanza-aprendizaje. Generar, analizar y comprender las configuraciones de entornos para la enseñanza y aprendizaje en línea implica, necesariamente, reconocer su enorme complejidad intrínseca, asociada a la gama de usos de estas tecnologías, a su diversidad y a la heterogeneidad de criterios utilizados

para describirlos y clasificarlos. Algunos criterios que suelen aparecer en las clasificaciones, bien de manera independiente o combinados, son los siguientes:

En primer lugar, la configuración de recursos tecnológicos utilizados: computadoras, redes más o menos amplias de computadoras, sistemas de interconexión, soporte y formato de la información, plataformas, sistemas de administración de contenidos o de aprendizaje, aulas virtuales, etc.

En segundo lugar, el uso de aplicaciones y herramientas que permiten la combinación de recursos, como simulaciones, materia les multimedia, tableros electrónicos, correo electrónico, listas de correo, grupos de noticias, mensajería instantánea, videoconferencia interactiva, etc.

En tercer lugar, la mayor o menor amplitud y riqueza de las interacciones que las tecnologías seleccionadas posibilitan.

En cuarto lugar, el carácter sincrónico o asincrónico de las interacciones; En quinto y último lugar, las finalidades y objetivos educativos que se persiguen y las concepciones implícitas o explícitas del aprendizaje y de la enseñanza en las que se sustentan (Bustos y Coll, 2010).

El nivel de aprovechamiento de los entornos virtuales de enseñanza aprendizaje, se ven influenciados por la infraestructura tecnológica, nivel de acceso a Internet, el uso de herramientas de administración de contenidos y las estrategias digitales de los usuarios.

2.2.2. Comunicación

Internet se ha convertido en la plataforma para la comunicación, las herramientas tecnológicas desarrolladas para la comunicación permiten comunicarse alrededor del mundo de una forma más rápida y precisa, reduciendo costes y tiempo en el servicio de comunicación entre individuos o grupos de interés. Se puede destacar las siguientes aplicaciones dedicadas a las comunicaciones y relaciones sociales (Skype, Facebook, Twitter, YouTube, etc.).

Hoy en día todos los procesos y métodos de comunicación giran en torno a Internet, lo que ha generado que tanto la televisión, medios impresos, radio, telefonía, migren sus datos y servicios a software y páginas web orientados a brindar los mismos servicios y en muchos casos mejorados a sus usuarios de forma masiva haciendo uso de la red (Druetta, 2006).

Se ve como las empresas tanto en el ámbito local, nacional e internacional han trasladado sus servicios a herramientas tecnológicas, las mismas que les brindan mayor cobertura de sus servicios y por ende tener mayor competitividad empresarial. Para ello, han hecho uso de varias herramientas tecnológicas como: páginas web, correos electrónicos, sistemas en línea, de acuerdo a los niveles de necesidad de cada una de las empresas y según los servicios que deseen prestar.

Los dispositivos electrónicos con los cuales se puede acceder a Internet evolucionan también drásticamente, de tal manera que hoy en día se accesa a Internet en cualquier parte del mundo que cuente con la infraestructura instalada para poder conectarse, ya que todos los dispositivos actuales (teléfonos inteligentes, tabletas, computadoras portátiles y demás) tienen compatibilidad para conectarse a dispositivos de acceso a Internet.

2.2.3. Entretenimiento

Como parte del ocio el Internet es una fuente que brinda muchas aplicaciones dedicadas al entrenamiento de los usuarios en varias secciones como: video, música, chat, juegos en línea, tv en línea, radio en línea, etc., Internet no solo se lo aprovecha en torno a acceder a información, sino también para entretenerse (relación con amigos, juegos en línea, películas, etc.), dependiendo los gustos de cada individuo.

2.3. Rendimiento académico

El rendimiento académico, se lo define como el nivel de éxito de un estudiante en materia de educación, este nivel de éxito puede variar de acuerdo a varios factores que inciden directa o indirectamente. Entre los factores determinantes del éxito o fracaso en el rendimiento académico se encuentran los determinantes personales (entre los más importantes se tiene las competencias cognitivas, la motivación, la inteligencia emocional, destrezas y habilidades del estudiante, etc.); determinantes sociales (entre los cuales se destaca los niveles sociales, culturales, de ingresos y el nivel de educación de los padres); y determinantes institucionales (los cuales pueden ser el ambiente institucional, condicionamientos institucionales, apoyo y orientación al estudiante) (Garbanzo, 2007).

La correlación de todos los determinantes mencionados anteriormente, influyen en el nivel de rendimiento académico de un estudiante, como se puede apreciar en la siguiente ilustración.

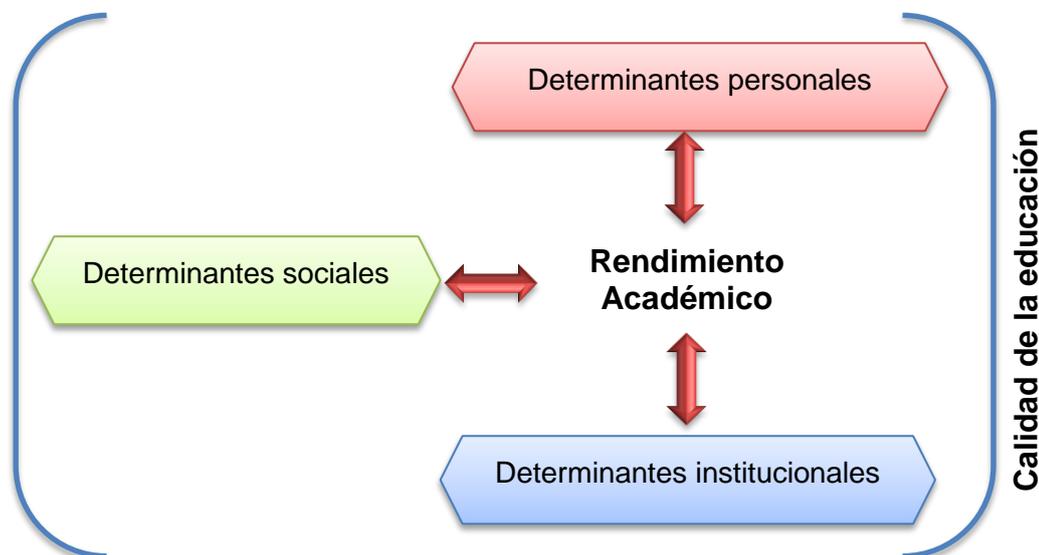


Figura 1: Determinantes del rendimiento académico
Fuente: (Garbanzo, 2007)

Por lo tanto, el éxito o fracaso en el rendimiento académico está estrechamente relacionado con las capacidades intelectuales del estudiante, la sociedad en la cual se desarrolla y las oportunidades tanto académicas y estratégicas brindadas por la institución educativa.

A la hora de delimitar qué factores están incidiendo en el éxito o fracaso escolar, no es extraño encontrarse con muchas dificultades, debido a que dichos factores o variables constituyen una red fuertemente entrelazada, por lo que resulta difícil delimitarlas para atribuir efectos claramente discernibles a cada una de ellas. En los últimos años, se ha dado un avance muy relevante en el sentido de superar los enfoques clásicos con una finalidad predictiva, pasando a otros más complejos con una finalidad explicativa a través de modelos estructurales o causales. Al conjunto de variables que indican en el éxito o fracaso se les conoce como condicionantes del rendimiento académico (Gonzalez, 2003).

El rendimiento académico está determinado por varios factores o variables, los cuales determinan de una u otra manera el grado de incidencia ya sea en el éxito o fracaso estudiantil. Entre los factores más importantes, que influyen en el rendimiento académico como estudiantes se puede enumerar: la automotivación, el esfuerzo realizado, el nivel cognitivo y haber escogido un área de estudio de su interés. Así mismo, los ingresos económicos de un individuo, pueden determinar si el estudiante tiene mayor facilidad de

acceso a recursos y materiales tecnológicos adecuados, para incrementar su nivel de competencia estudiantil.

2.4. Minería de datos

En la era de la información, todas las empresas han incorporado sistemas de información, los que generan gran cantidad de datos, y que son utilizados para extraer conocimiento a partir de ellos.

Para Hernández, Ramírez y Ferri (2007), los datos dejan de ser el resultado histórico de los sistemas de información, para ser la materia prima, la cual debe ser explotada para obtener el conocimiento, que será aplicado en la toma de decisiones de las empresas.

Pérez y Santin (2008), expresan que la minería de datos, permite generar conocimiento a partir de datos, mediante la aplicación de técnicas y herramientas las mismas que tienen como finalidad la extracción de conocimiento. Por lo tanto, se convierte en la fase fundamental del proceso de extracción de conocimientos a partir de los datos (KDD).

2.4.1. Tipos de datos

La minería de datos se puede aplicar a cualquier tipo de información. Pero tanto las técnicas como herramientas que se aplican a las bases de datos difiere dependiendo el tipo de datos a los cuales se aplica. Hernández, Ramírez y Ferri (2007) presentan la siguiente clasificación:

- **Bases de datos relacionales:** se trata de una compilación de relaciones (tablas) las cuales están conformadas por atributos (columnas o campos) y a su vez tienen un número de tuplas (filas o registros). La tupla representa a un objeto y tiene una clave única o primaria la cual la identifica.
- **Bases de datos espaciales:** este tipo de bases de datos están constituidas por información relacionada con el espacio físico que puede incluir datos geográficos, imágenes médicas, redes de transporte, etc. dentro de este tipo de datos la minería de datos permite obtener patrones entre los datos.
- **Bases de datos temporales:** están conformados con atributos relacionados con el tiempo. La minería de datos aplicada a este tipo de datos permite encontrar características de tendencia o evolución de los datos.
- **Bases de datos documentales:** almacena las descripciones de los objetos, éstas pueden ser documentos estructurados, semi-estructurados y no estructurados. Al aplicar

minería de datos se puede obtener asociación, agrupación o clasificación de los objetos textuales.

- **Bases de datos multimedia:** puede estar conformada por datos multimedia (audio, imágenes y vídeo), son bases de datos que soportan objetos con gran cantidad de tamaño.
- **Word Wide Web:** la web en la actualidad es el almacén con mayor cantidad de datos, el reto de la minería de datos es poder extraer el conocimiento de ella, ya que no todos los datos almacenados en la web cumplen un estándar sobre su estructura.

2.4.2. Tipos de modelos

El principal objetivo de la minería de datos es la extracción del conocimiento, este tipo de conocimiento se lo puede generar de diferentes maneras. Los tipos de modelos pueden ser: predictivos y descriptivos según Hernández, Ramírez y Ferri (2007).

- **Modelos predictivos:** evalúan valores futuros o desconocidos de variables, usando otras variables o campos de la base de datos como variables independientes o predictivas. Cuyos resultados son denominados variables objetivo o dependiente.
- **Modelos descriptivos:** se utilizan para explorar propiedades de los datos examinados a través de patrones que los explican y resumen.

2.4.3. Fases del proceso de minería de datos

El proceso de extracción de conocimiento es interactivo e iterativo; iterativo es porque se debe de realizar varias iteraciones para poder extraer conocimiento de buena calidad; así mismos, debe de ser interactivo en vista de que deben de estar inmersos en el proceso de preparación de los datos especialistas o experto en el problema, para garantizar la calidad del conocimiento extraído.

Entre las fases del proceso de la minería de datos Moreno (2007) detalla:

Integración y recopilación

En la primera fase del proceso de minería de datos, se debe identificar la fuente de datos, así como la recopilación e integración de los mismos en un repositorio o base de datos. Los datos pueden originarse tanto interna como externamente a la organización. Esta fase es

esencial para determinar la calidad de los datos recopilados y por ende garantizar la extracción de conocimiento de calidad.

Selección, limpieza y transformación

La calidad del conocimiento extraído se ve afectada por la calidad de los datos, por lo tanto, la fase de selección, limpieza y transformación tiene como objetivo determinar los datos que no están dentro del comportamiento normal del conjunto de datos (valores atípicos o poco comunes), así como, determinar y encontrar los valores perdidos.

En esta fase del proceso de minería de datos se busca eliminar los datos irrelevantes, los cuales no ofrecen información valiosa para la extracción de conocimiento; también, determinar si se encuentran datos perdidos, lo cual suele ser provocado por errores en la herramienta o dispositivo de recolección de datos. La presente fase determina el grado de relevancia del conocimiento, ya que a partir de esta fase se determina la calidad de los datos y por ende la calidad del conocimiento extraído.

Minería de datos

Es la fase que caracteriza al proceso, ya que en ella se genera conocimiento el cual será de gran utilidad para el usuario. La extracción o generación de conocimiento se lo hace mediante la aplicación de modelos. Los modelos ayudan a describir los patrones y relaciones existentes entre los datos.

Para el correcto desarrollo de la presente fase se debe partir de determinar cuál tarea de la minería de datos se adapta y es la más apropiada; también seleccionar el tipo de modelo a utilizar y por último determinar el tipo de algoritmo que resuelva la tarea y permita obtener el modelo buscado.

Entre las tareas de la minería de datos Moreno (2007) describe:

Clasificación: es una tarea predictiva muy utilizada. Permite mejorar el nivel de precisión de las nuevas instancias. Cada instancia pertenece a una clase, la cual está determinada por el valor de un atributo, el cual se lo conoce como clase o instancia. Los atributos relevantes de la instancia se usan para determinar o predecir la clase.

Regresión: permite aprender una función real que asigna a cada instancia un valor real. El valor a predecir de esta tarea es numérico. El objeto de esta tarea es reducir el error entre el valor real y el valor predicho.

Agrupamiento o clustering: se trata de una tarea descriptiva, y se la usa para obtener grupos a partir de los datos. Los datos son agrupados basándose en el principio de aumentar la similitud o igualdad entre un grupo y reduciendo la similitud o igualdad entre dos grupos distintos.

El algoritmo mayormente utilizado para realizar esta tarea es K-means, por ser eficaz y veloz. Este algoritmo divide los datos en k subconjuntos no vacíos, usando algún tipo de heurística o aleatoriamente; seguidamente calcula el centro (centroide) de cada partición el cual es llamado el punto medio del clúster, con esto procede a ir asignando o sumando cada dato al clúster cuyo centro tenga mayor grado de similitud. Este proceso se repite, hasta obtener la afinidad total de los datos.

Correlación o factorización: es la tarea descriptiva que se la utiliza para determinar el grado de igualdad o similitud de los valores de dos variables numéricas. El objeto de esta tarea es determinar la relevancia de atributos, encontrar atributos repetidos y dependencia entre ellos. El resultado de esta tarea es conocer la correlación lineal o relaciones existentes entre variables.

CAPITULO III
METODOLOGÍA

En el presente capítulo, se describe la metodología, mediante la cual se ha desarrollado el presente trabajo, en lo cual se detalla cada uno de los métodos estadísticos y de análisis que se aplicaron a los datos, entre las técnicas y procesos para comprobar las hipótesis de la investigación se utilizó: análisis clúster o de conglomerados, análisis discriminante, varianza y regresión logística binomial.

3.1. Población y muestra

La población objeto de estudio, son los estudiantes de la Universidad Metropolitana de Quito a partir del segundo ciclo.

El presente proyecto tiene una muestra finita, por tal motivo se aplicó la fórmula de muestra para poblaciones finitas, como recomienda Herrera (2009) el cual presenta la siguiente denotación:

$$n = \frac{N * Z_{\alpha}^2 * p * q}{d^2 * (N - 1) + Z_{\alpha}^2 * p * q}$$

De dónde:

N = Total de la población

Z α = Valor del porcentaje de nivel de seguridad

p = proporción esperada

q = 1 – p

d = Nivel de precisión

Por lo tanto, reemplazando valores en la fórmula sugerida por Herrera (2009), teniendo una población de 1400 alumnos (N=1400); como porcentaje de seguridad el 95% (Z α = 1,96); como valor de la proporción esperada el 5% (p=0,5), q = 1-0,5 =0,5 y como nivel de precisión el 5% (0,05), se obtiene:

$$n = \frac{1400 * 1,96^2 * 0,5 * 0,5}{0,05^2 * (1400 - 1) + 1,96^2 * 0,5 * 0,05}$$

$$n = \frac{1400 * 3,84 * 0,5 * 0,5}{0,0025 * 1399 + 3,84 * 0,5 * 0,05}$$

$$n = \frac{1344}{3,4975 + 0,096}$$

$$n = \frac{1344}{3,5935}$$

$$n = 374,0086 = \mathbf{374}$$

Aplicando la fórmula, la muestra es de 374 siguiendo los requerimientos del proceso de investigación para el desarrollo del trabajo de investigación se aplicó un total de 560 encuestas.

3.2. Integración y recopilación

Para poder acceder a las instalaciones de la Universidad Metropolitana de Quito, se solicitó autorización al Ing. Jorge Luis Basantes Director Académico, el mismo que autorizó la aplicación de las encuestas para el levantamiento de información para el presente proyecto. Ver Anexo 33.

Para la recolección de datos se hizo uso de encuestas dirigidas a estudiantes y de una entrevista realizada al Jefe del Departamento Informático de la Universidad Metropolitana.

La encuesta se aplicó a los alumnos de la Universidad Metropolitana de Quito, a partir del segundo ciclo en adelante en las diferentes carreras ofertadas por la Institución. La encuesta fue elaborada por el Director General del Proyecto Dr. Juan Carlos Torres Díaz (Ver Anexo 1) se encuentra estructurada para poder recolectar información referente al estudiante, la carrera, información demográfica (edad, género e ingresos económicos de la familia), en torno al acceso a Internet se indagó acerca de: lugar habitual de conexión, días de conexión a la semana, nivel de conocimiento y los años de experiencia en acceso a Internet. También se pregunta acerca del uso que se le da al Internet tanto para temas académicos como para entretenimiento. Para recolectar información acerca del nivel de rendimiento académico del alumno se consultó acerca de las materias tomadas y aprobadas en el nivel anterior.

La encuesta está conformada por 17 literales, los que están agrupados de acuerdo al tipo de información a extraer.

La primera sección de la encuesta está orientada a recabar información acerca de la universidad, carrera del estudiante; así como, también los datos demográficos del alumno

(edad, género, ingresos familiares); también temas acerca del lugar habitual de acceso a Internet, días de conexión a la semana, niveles de conocimiento en el uso de Internet, horas de conexión diaria y grado de experiencia en el uso de Internet; así mismo, se establecen cuestiones que tienen como objeto levantar información entorno al uso de Internet tanto en el ámbito académico y de entretenimiento, para conocer el nivel de rendimiento académico se cuestionó sobre el número de materias matriculadas y el número de materias aprobadas en el ciclo anterior.

Las encuestas¹ se aplicaron a los alumnos en las instalaciones físicas de los campus universitarios de la Universidad Metropolitana de Quito previa la autorización de la máxima autoridad de dicha institución educativa, para su posterior ingreso a la plataforma web SurveyMonkey, en el link <https://es.surveymonkey.com/r/encuestaUMetropolitana> el cual fue desarrollado por el director del proyecto. Seguidamente se obtuvo una base de datos con un total de 560 registros.

En lo concerniente a la entrevista², aplicada al Jefe del Departamento Informático de la Universidad Metropolitana Ing. Héctor Hernández, las preguntas de la entrevista fueron orientadas a recabar información concerniente a dos parámetros:

➤ **Instalaciones física**

- Infraestructura de las salas de cómputo para estudiantes.
- Tipos de conexiones de Internet.
- Capacidad de ancho de banda.
- Cantidad de estudiantes.
- Software para la gestión administrativa y académica.
- Entorno virtual de aprendizaje.
- Servicios de la plataforma virtual.
- Cantidad de alumnos con computadores personales.

➤ **Políticas de uso de tecnología**

- Las obligaciones de los docentes en el uso de elementos tecnológicos en los procesos de enseñanza.
- Planes de formaciones de docentes orientados al uso de tecnologías para la educación.
- Nivel de capacitación de los profesores en temas tecnológicos.
- Nivel de capacitación de los profesores en temas pedagógicos.

¹ **Encuesta a Estudiantes UMET**, Disponible en: <https://www.youtube.com/watch?v=zQ5ZmIS-Blg&index=3&list=LL2YIUpw-76WGTB-CH5o4BUQ>

² **Entrevista Docente UMET**, Disponible en: <https://www.youtube.com/watch?v=R5ivyqeDlyU&list=LL2YIUpw-76WGTB-CH5o4BUQ&index=2>

3.3. Selección, limpieza y transformación

Luego del proceso de recolección de datos, en los cuales se aplicó 560 encuestas a los alumnos de la Universidad Metropolitana de Quito, se procedió a realizar el respectivo proceso de selección, limpieza y transformación de los datos.

La tarea de limpieza de los datos contempla la eliminación de archivos con datos atípicos, irrelevantes y datos en blanco, con lo cual se redujo el número de registros a 470, con esto se obtuvo una base de datos de calidad, para su posterior procesamiento y análisis.

En lo referente a datos vacíos o en blanco, para no eliminar el registro por su nivel de similitud con la tendencia de la muestra, se procedió a reemplazar los datos vacíos con la moda de la variable pertinente.

De igual manera en esta fase se procedió a realizar la transformación de variables y creación de nuevas variables, en lo referente a la transformación de variables se tomó como variables a transformar: el sexo, los niveles de ingresos, lugar de conexión (Tabla 1); y para la creación de nuevas variables se utilizó como variables de origen el número de materias matriculadas y el número de materias aprobadas en el semestre anterior obteniendo como resultado la variable de rendimiento académico del alumno (Tabla 2).

Tabla 1: Transformación de variables

Variable a intervenir		Operación realizada	Variable resultante	
Nombre	Valores		Nombre	Valores
Sex	Hombre	Numerización	Sex	1
	Mujer			2
Ing	Hasta 350 dólares	Numerización	Ing	Ingreso 1
	Hasta 600 dólares			Ingreso 2
	Hasta 1000 dólares			Ingreso 3
	Hasta 1500 dólares			Ingreso 4
	Más de 1500 dólares			Ingreso 5
Lug_con	Desde la casa	Numerización	Lug_con	Lugar 1
	Desde un cyber café			Lugar 2
	Desde el trabajo			Lugar 3
	Desde la Universidad			Lugar 4
	Red Móvil			Lugar 5

Fuente: *Elaboración propia.*

Tabla 2: Generación de nueva variable rendimiento académico

Variable a intervenir		Operación realizada	Variable resultante	
Nombre	Valores		Nombre	Valores
sem_ant_asi_mat	Materias matriculado	Discretización	ren_aca	Aprobado
sem_ant_asi_apr	Materias aprobadas			Reprobado

Fuente: *Elaboración propia.*

Seguidamente del proceso de limpieza y transformación de variables, se realizó la determinación de las variables tanto dependientes como independientes, las cuales serán intervenidas para la comprobación de las diferentes hipótesis planteadas en el proyecto, como se puede ver en la Tabla 3.

Tabla 3: Determinación de variables para comprobación de hipótesis

Grupo inicial	Variables	Variable final (Perfil)	Hipótesis de aplicación	Tipo
Ingresos familiares	Ingreso 1	Nivel de ingreso	Hipótesis 1	Independiente
	Ingreso 2			
	Ingreso 3		Hipótesis 2	
	Ingreso 4			
	Ingreso 5			
Académico	a_ing_pla	Perfil académico	Hipótesis 1	dependiente
	a_con_pro			
	a_con_est			
	a_rec_edu			
	a_vid_aca			
	a_for_vir		Hipótesis 3	Independiente
	a_pos_aca			
	a_hor_cha			
	a_hor_inf			
	a_hor_bib			
Entretenimiento o diversión	e_cha_div	Perfil de entretenimiento	Hipótesis 2	Dependiente
	e_red_soc			
	e_jue_lin		Hipótesis 4	Independiente
	e_mus_vid			
	e_vid_you			
Nivel académico	sem_ant_así_mat	Rendimiento académico	Hipótesis 3	Dependiente
	sem_ant_así_apr		Hipótesis 4	

Fuente: *Elaboración propia.*

3.4. Minería de datos

El proceso de minería de datos permite extraer el conocimiento a partir de los datos, el cual debe de ser entendible y útil para el usuario, en la presente fase se determinó las relaciones existentes entre variables, así mismos se agrupó las variables de acuerdo a su nivel de similitud para determinar los perfiles académicos y de entretenimiento. De igual manera, se procedió a comprobar las hipótesis.

3.5. Tarea de minería de datos a utilizar: clusterización

Se puede definir al análisis clúster (conglomerados), como el proceso para determinar o clasificar grupos, la cual se basa en la técnica de análisis exploratorios de datos como expresa Villardón (2011). Con este análisis se busca clasificar los grupos, para poder determinar su diferenciación y conformación de cada uno, de acuerdo a variables medidas dentro de cada grupo Andes (2009). De tal manera, que con el análisis clúster se logra determinar grupos, los cuales presentan similitudes entre si y diferencias con otros grupos (Voges, 2009).

Existen dos tipos de algoritmos para la determinación de conglomerados (De la Fuente, s.f.):

- **Algoritmos de partición:** en esta forma de clasificar, el usuario previamente determina el número de clúster o conglomerados que desea clasificar.
- **Algoritmos jerárquicos:** este tipo de algoritmo a su vez se clasifica en:
 - **Aglomerativo:** consiste en que cada observación va formando un conglomerado, y seguidamente las observaciones que concuerdan o están cerca se van uniendo y terminan formando conglomerados con diferencias establecidas en base a la medición de sus variables.
 - **Disociativo:** su proceso en cambio parte de un conglomerado general, el mismo que se va dividiendo en observaciones según criterios y al final se generan los clúster de acuerdo a los coeficientes de sus variables.

Para grupos grandes de datos, la mayoría de autores citados recomiendan la aplicación de conglomerados no jerárquicos, es decir, de partición, de tal manera que se establece en primera estancia el número de conglomerados que se desea obtener, en base a las variables de medición.

El procedimiento k-medias, es el método más sencillo de clasificación, para el cual primeramente se debe de establecer el número de clúster que desea obtener del proceso, conocido como parámetro (k), el proceso determina los nuevos centroides para cada clúster, seguidamente se repite para todas las observaciones y se van asignando a los centroides con mayor similitud (Molina y Herrero, 2006). Este procedimiento debe obedecer a 3 reglas básicas (Martínez, s.f.) :

- 1) Cada elemento (observación o variable) pertenece a uno y sólo uno de los grupos.
- 2) Todo elemento debe quedar clasificado.
- 3) Cada grupo debe ser internamente similar.

Para realizar el proceso de k-medias, se utiliza el siguiente algoritmo (Nadu, 2011):

- Asignar valores iniciales para medias $m_1, m_2 \dots m_n$.
- Asignar a cada elemento de la agrupación que tiene centro más cercano.
- Calcular el nuevo centro para cada clúster hasta que se cumpla el criterio de convergencia.

De tal manera que según Nadu (2011); el primer paso para realizar el procedimiento de k-medias es, asignar valores de iniciación para cada clúster, luego a cada observación o variable se la va asignando al centroide más cercano. Por último, se recalcula los centros de los clúster, para determinar el grado de concordancia interna en cada grupo.

3.6. Análisis discriminante

El análisis discriminante es una técnica de análisis multivariado, permite clasificar las observaciones de análisis, en grupos determinados a priori, además permite analizar las causas que determinan o forman a cada grupo establecido (Ramírez y Castillo, 2002) (Tusell, 2012). El análisis discriminante permite comprobar el porcentaje de clasificación de cada grupo, y ayuda a determinar la clasificación más idónea para el análisis y discusión de los resultados.

3.7. Comprobación de hipótesis: regresión logística

En vista de que las variables dependientes son de tipo binario o dicotómica, se procede a realizar la comprobación de hipótesis mediante regresión logística binomial, la cual permitirá predecir la probabilidad de los diferentes resultados para cada análisis.

Para dar por cumplida una hipótesis mediante regresión binomial, debe de cumplir ciertos parámetros (Torres, 2012):

Prueba de ómnibus: ayuda a comprobar que mínimo uno de los coeficientes que conforman el modelo sean distintos a cero (0), y para poder aceptar el modelo debe presentar un valor significativo ($p < 0,05$).

Prueba de Hosmer Lemeshow: permite calcular la bondad de ajuste del modelo de regresión, que calcula para cada observación el grupo de datos de las probabilidades de la

variable dependiente a predecir. Para poder concluir que el modelo se ajusta a los datos ($p > 0,05$).

Bondad de ajuste: se lo utiliza para determinar si el modelo se ajusta a los datos, el valor de significancia debe ser menor a 0,05 ($p < 0,05$), los valores de análisis son arrojados por la prueba de verosimilitud (Chi-cuadrado) y Pearson y Deviance.

R²Nagelkerke: permite conocer el porcentaje de exactitud de un modelo, muestra la varianza descrita o explicada por el modelo; para poder aceptar el modelo debe de describir al menos el 70% de los datos.

Test de Wald: se lo utiliza para ajustar el modelo o coeficiente de determinación, permite verificar la hipótesis nula ($\beta_i=0$), la significación estadística asociada y el valor OR ($\exp(B)$) con sus intervalos de confianza.

CAPÍTULO IV
RESULTADOS

4.1. Descripción de la muestra

A continuación se muestra los resultados obtenidos de las encuestas a los 470 estudiantes de la Universidad Metropolitana de Quito, los mismos que en lo posterior servirán para comprobar las hipótesis planteadas.

De una población de 470 alumnos encuestados de las diferentes carreras que oferta la Universidad Metropolitana de Quito a partir del segundo semestre de estudio, 221 pertenecen al género femenino que corresponde al 47% del total de la muestra y 249 pertenecen al género masculino, que corresponde al 53%. Tal como lo muestra la Figura 2.

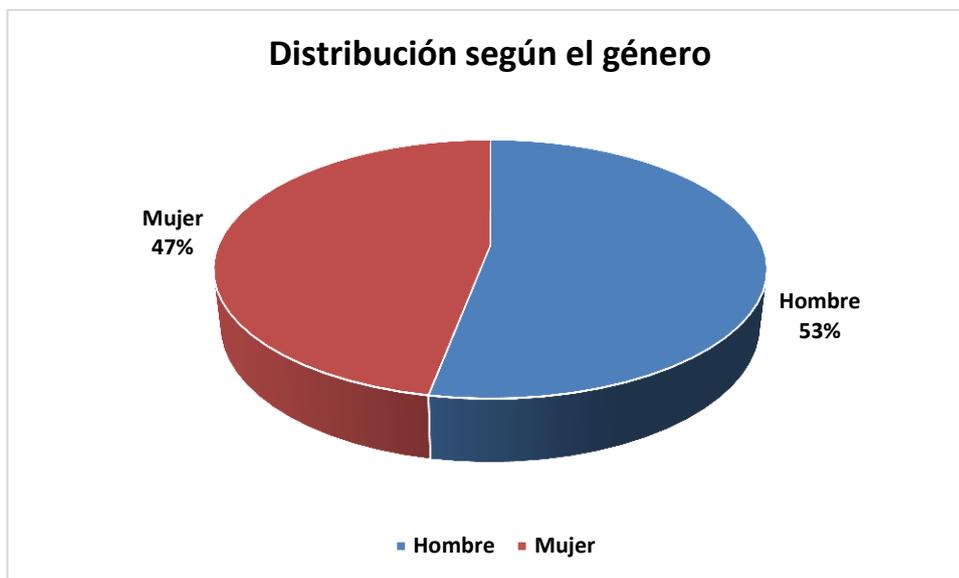


Figura 2: Distribución de los alumnos por género
Fuente: Elaboración propia

Actualmente la Universidad Metropolitana de Quito, oferta 8 carreras, las cuales tienen los siguientes porcentajes de alumnos: derecho 24,7%, optometría 19,6%, gestión empresarial 15,3%, gestión de empresas turísticas y hoteleras 11,5%, ingeniería en sistemas de información 10,2%, ciencias administrativas y contables – CPA 7,9%, diseño gráfico 7,4% y talento infantil con 3,4%, como se muestra en la Figura 3.

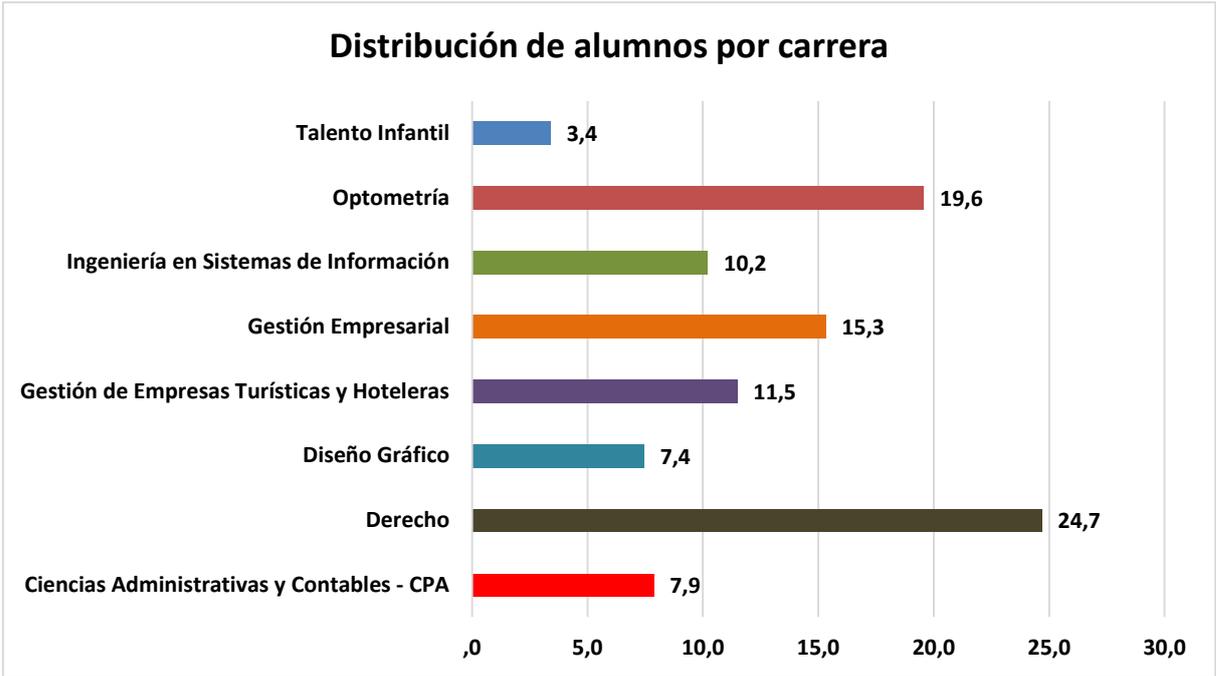


Figura 3: Distribución de alumnos por carrera
Fuente: Elaboración propia

Las edades de los encuestados se encuentran dispersas entre 18 y 35 años de edad, existiendo una mayor concentración de alumnos con 19 años de edad con un 11,5%, seguidos de los alumnos con 21 años de edad con un 10,4% y el porcentaje sobrante se encuentra distribuido entre las edades restantes. Como se lo puede observar en la Figura 4.

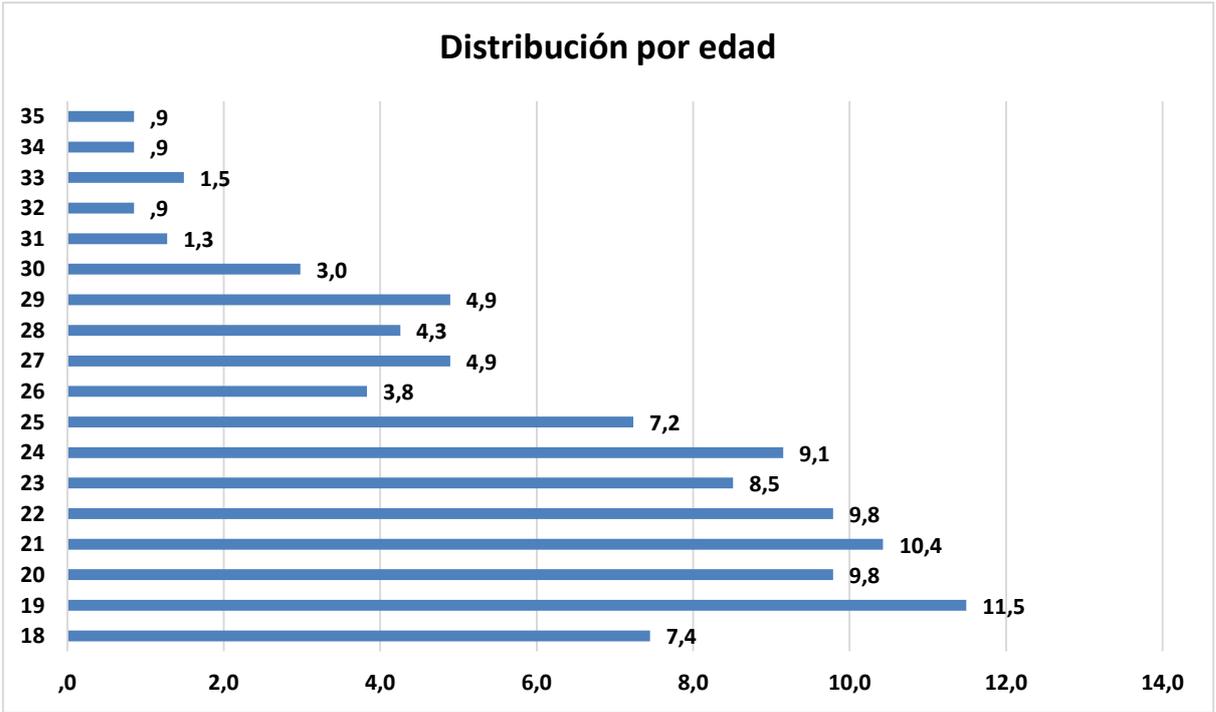


Figura 4: Distribución por edades
Fuente: Elaboración propia

La distribución de los ingresos mensuales de las familias de los estudiantes se clasifica de la siguiente manera: el 23,8% está concentrado en familias con ingresos de hasta 1000 dólares, 23,6% tienen ingresos familiares de más de 1500 dólares, 20,6% ingresos de hasta 600 dólares, 19,1 hasta 1500 dólares y el 12,8% restante tienen ingresos de 350 correspondiente a la salario básico, como se puede observar en la Figura 5.

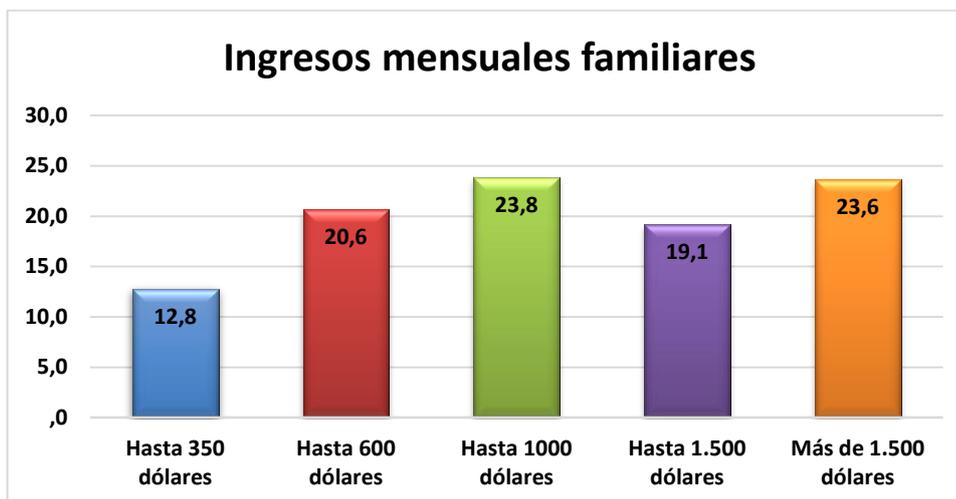


Figura 5: Ingresos mensuales de las familias de los alumnos
Fuente: Elaboración propia

4.2. Acceso a Internet: lugar de conexión, tiempo y nivel de conocimiento

Los lugares de conexión habitual de los alumnos a Internet están distribuidos de la siguiente manera: el 65,1% se conecta desde su casa, 17,4% se conectan de algún tipo de red móvil, el 11,7% desde su lugar de trabajo, el 3% desde un cyber y el 2,8% restante desde la Universidad a la cual asisten como se puede visualizar en la Figura 6.

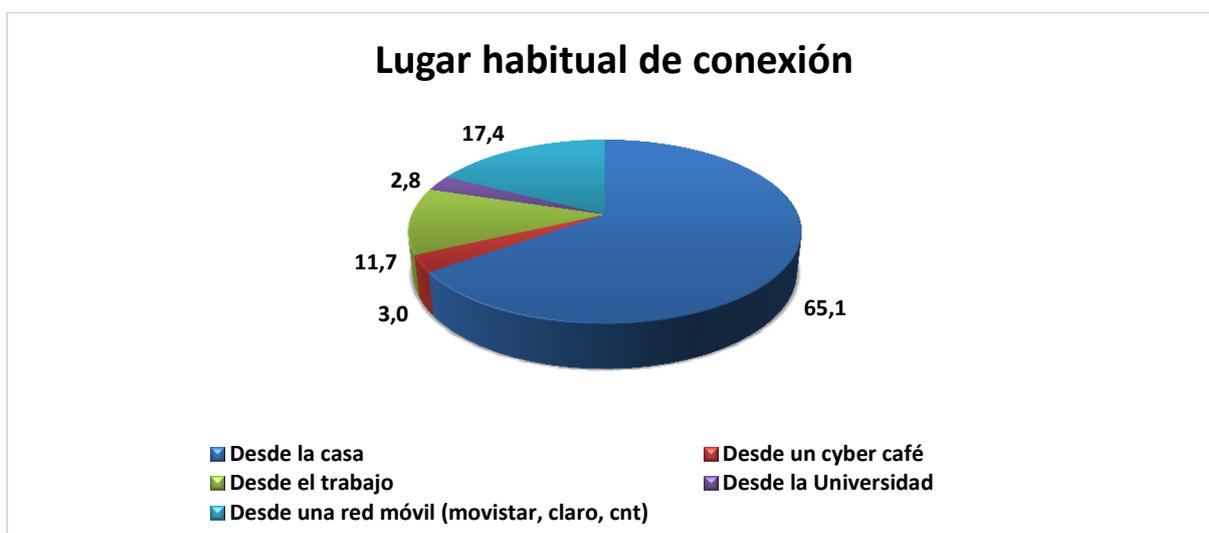


Figura 6: Distribución de los alumnos según el lugar de conexión
Fuente: Elaboración propia

En lo referente a cuantos días a la semana se conectan a Internet se puede observar en la Figura 7, que el mayor porcentaje de alumnos se conectan 7 días a la semana con un 66%, mientras que en menor porcentaje un 0,4% se conectan únicamente una vez por semana.

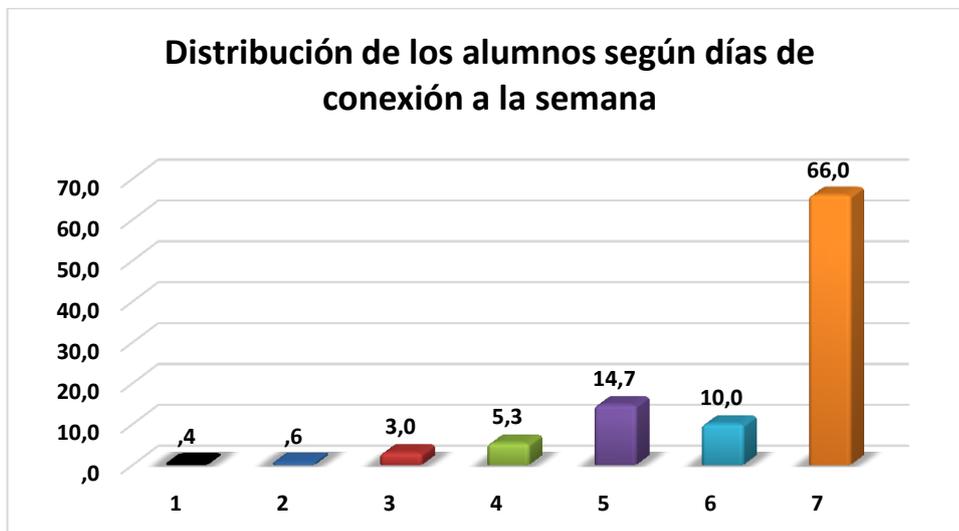


Figura 7: Días de conexión a la semana
Fuente: Elaboración propia

El nivel de conocimiento en el uso de Internet de los estudiantes esta medido en un rango del 1 al 10. De los alumnos encuestados se puede ver que el 31,1% se centra en un nivel 8, el 39,1% se distribuye entre alumnos con niveles de conocimientos de entre 1 y 7, mientras que el 29,8% se encuentra disperso entre nivel 9 y 10, como se puede observar en la Figura 8.

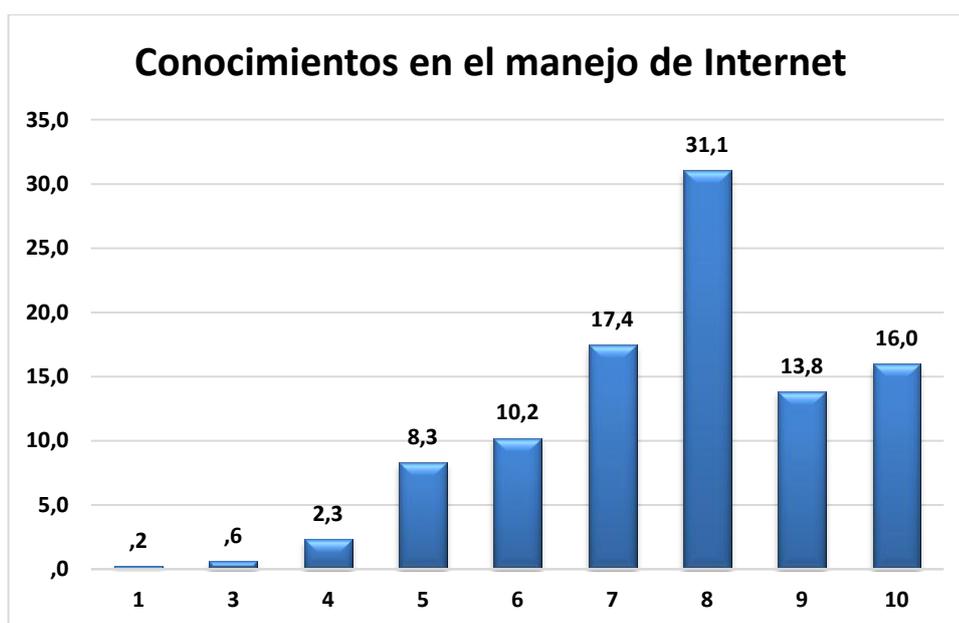


Figura 8: Nivel de conocimiento en el manejo de Internet
Fuente: Elaboración propia

Las horas de conexión de los alumnos se encuentran con mayor concentración entre 4 y 6 horas diarias con un porcentaje del 38,30, seguido de los alumnos entre 1 y 3 horas diarias con un 28,72 y el menor grado de concentración en los alumnos que se conectan de 13 a 15 horas con un 3,83, como se puede observar en la Figura 9.

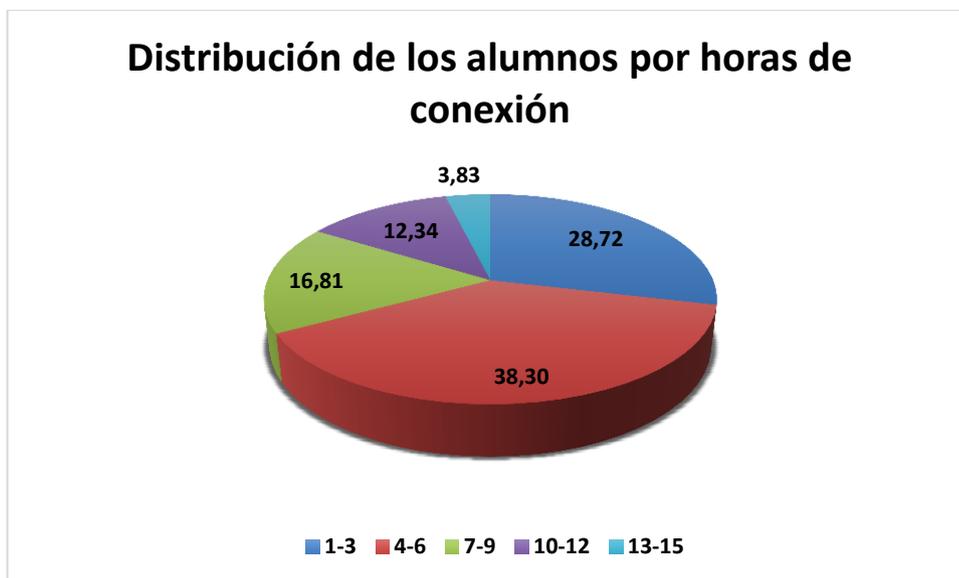


Figura 9: Distribución de los alumnos por horas de conexión
Fuente: Elaboración propia

En cuanto a los años de experiencia en el uso de Internet, los alumnos se encuentran con mayor densidad de 5 a 6 años con un 35,96 y en menor densidad en alumnos con 16 a 18 años con un porcentaje del 0,85; como se puede visualizar en la Figura 10.

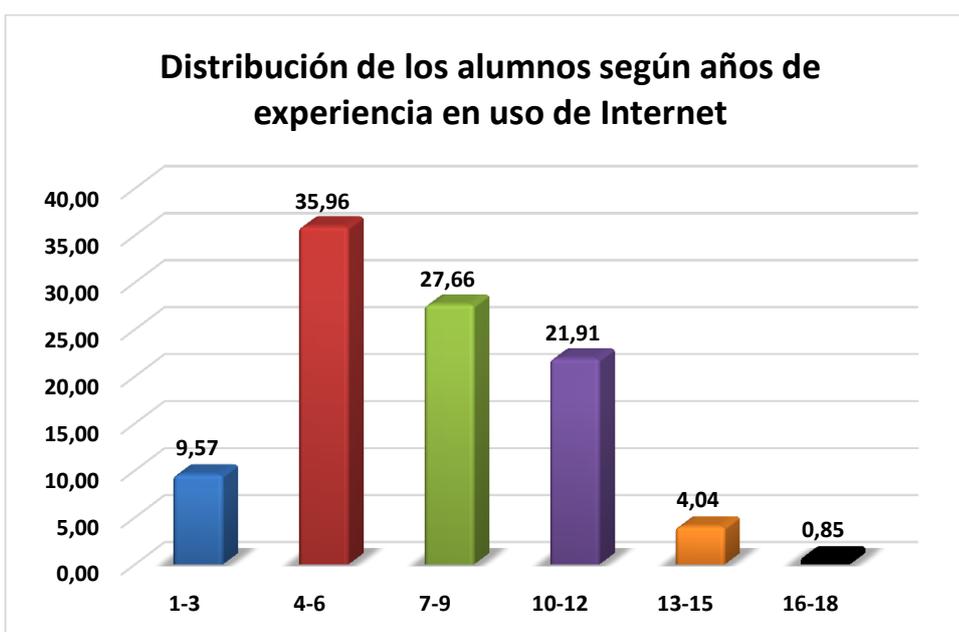


Figura 10: Años de experiencia en el uso de Internet
Fuente: Elaboración propia

4.3. Usos de Internet en el ámbito académico y de entretenimiento

4.3.1. Usos de Internet en el ámbito académico

En lo referente al uso de Internet en el ámbito académico, se puede observar en la Tabla 2 que los alumnos en un porcentaje del 33,6% se conectan mínimo una vez a la semana a la plataforma virtual, seguido de los alumnos que se conectan dos veces a la semana con un 22,1%; y el 18,3% se conecta entre 3 y 15 veces semanales; además, se puede observar que un 13,6% no se conectan a la plataforma.

Tabla 4: Ingresos semanales a la plataforma virtual

Ingresos semanales	Frecuencia	Porcentaje
0,00	64	13,62
1,00	158	33,62
2,00	104	22,13
3,00	58	12,34
4,00	27	5,74
5,00	31	6,60
6,00	4	0,85
7,00	6	1,28
8,00	3	0,64
9,00	1	0,21
10,00	11	2,34
11,00	1	0,21
15,00	2	0,43
Total	470	100,00

Fuente: Elaboración propia

Como se puede observar en la Tabla 3, el 16,81% de los alumnos no realizan consultas al docente a través de la plataforma virtual, un 13,62% de los alumnos hacen dos consultas mensuales al docente, un 9,15% realiza al menos una consulta al mes, mientras que el porcentaje restante se distribuye entre 3 a 10 preguntas mensuales al docente.

Tabla 5: Consultas mensuales al profesor

Consultas al Docente	Frecuencia	Porcentaje
0,00	79	16,81
1,00	43	9,15
2,00	64	13,62
3,00	44	9,36
4,00	36	7,66
5,00	53	11,28
6,00	17	3,62
7,00	12	2,55
8,00	10	2,13
9,00	5	1,06
10,00	57	12,13
11,00	3	0,64
12,00	4	0,85
13,00	1	0,21
14,00	1	0,21
15,00	9	1,91
16,00	2	0,43
18,00	2	0,43
19,00	2	0,43
20,00	20	4,26
30,00	6	1,28
Total	470	100,00

Fuente: Elaboración propia

Así mismo, en consultas a sus compañeros a través de la plataforma virtual se puede ver en la Tabla 4, que el 15,74% se aglomeran en que no realizan ninguna consulta; el 13,40% realiza 10 preguntas mensuales, el 11,49% realiza dos preguntas mensuales, mientras tanto el porcentaje restante se distribuye entre los demás valores.

Tabla 6: Consultas mensuales a compañeros

Consultas mensuales a compañeros	Frecuencia	Porcentaje
0,00	74	15,74
1,00	26	5,53
2,00	54	11,49
3,00	44	9,36
4,00	44	9,36
5,00	44	9,36
6,00	17	3,62
7,00	10	2,13
8,00	20	4,26
9,00	5	1,06
10,00	63	13,40
11,00	2	0,43
12,00	2	0,43
14,00	1	0,21
15,00	17	3,62
17,00	1	0,21
18,00	1	0,21
20,00	30	6,38
22,00	2	0,43
24,00	1	0,21
25,00	1	0,21
30,00	11	2,34
Total	470	100,00

Fuente: Elaboración propia

El número de descargas de recursos educativos desde la plataforma virtual de la universidad, se distribuye de la siguiente manera: con mayor concentración de alumnos que descargan entre 0 y 2 recursos educativos con un 58,09%; seguido de los alumnos con descargas de 3 a 5 recursos educativos con un 22,13%; así mismo, la concentración más baja de descargas de recursos educativos es de 12 a 14 con un 0,43%, como se puede ver en la Figura 11.

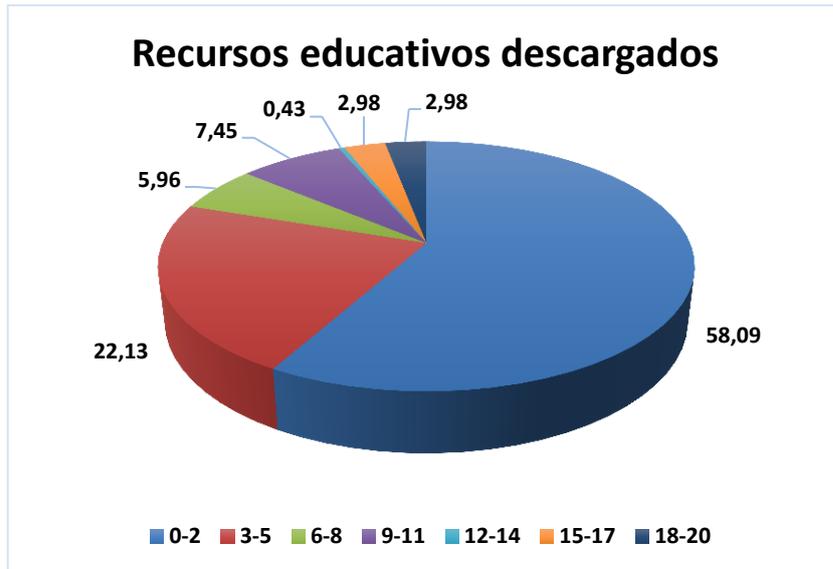


Figura 11: Descarga de recursos educativos de la plataforma virtual
Fuente: Elaboración propia

Los resultados indican que la mayoría de alumnos miran entre 0 a 2 videos de carácter académico en YouTube mensualmente con un 40,0%; seguido de los alumnos que observan entre 3 a 5 videos mensuales con un 32,77%; así mismo, se observa que la menor concentración de alumno observan entre 12 y 14 videos con un 0,85%; como se puede visualizar en la Figura 12.

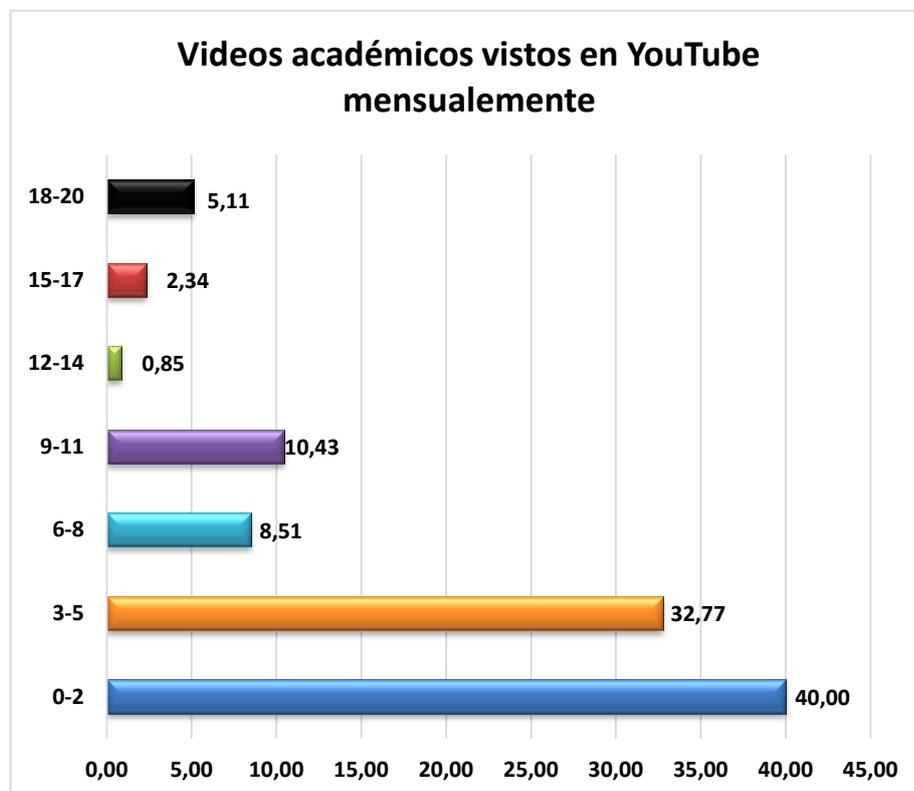


Figura 12: Videos académicos visualizados en YouTube mensualmente
Fuente: Elaboración propia

Se determina que la mayoría de alumnos no participan en foros virtuales con un 45,3%; mientras que el 44,7% restante se distribuye entre alumnos que interactúan entre 1 y 5 foros mensuales con una desviación estándar de 1,42, como se puede observar en la Figura 13.

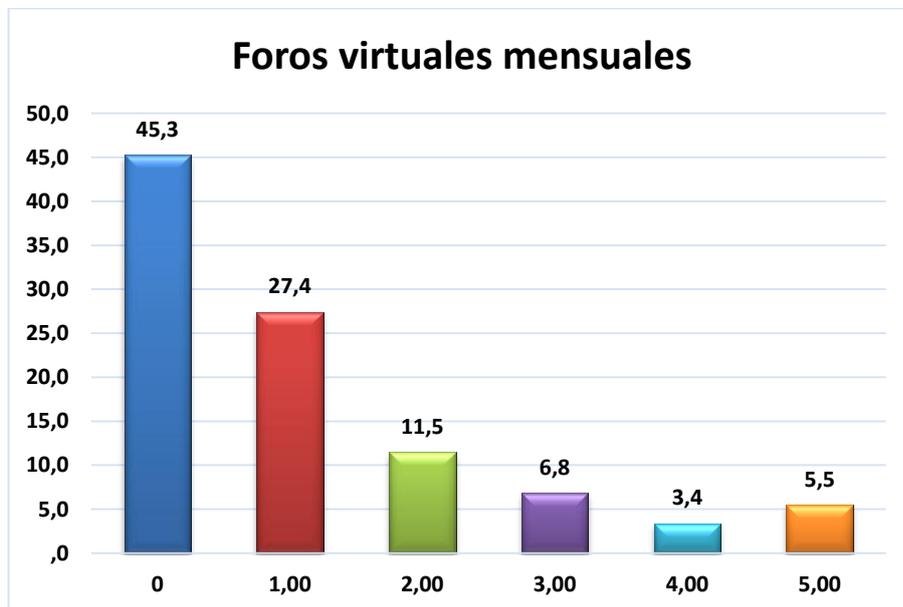


Figura 13: Foros virtuales mensuales
Fuente: Elaboración propia

De los resultados obtenidos, se determina que el 76,66% de los alumnos realizan entre 0 a 2 post/tweets mensuales de carácter académico en redes sociales, y el 32,34% se distribuye entre 3 y 20 post/tweets mensuales de índole académico en redes sociales con una distribución típica de 4,18; así como se muestra en la Figura 14.

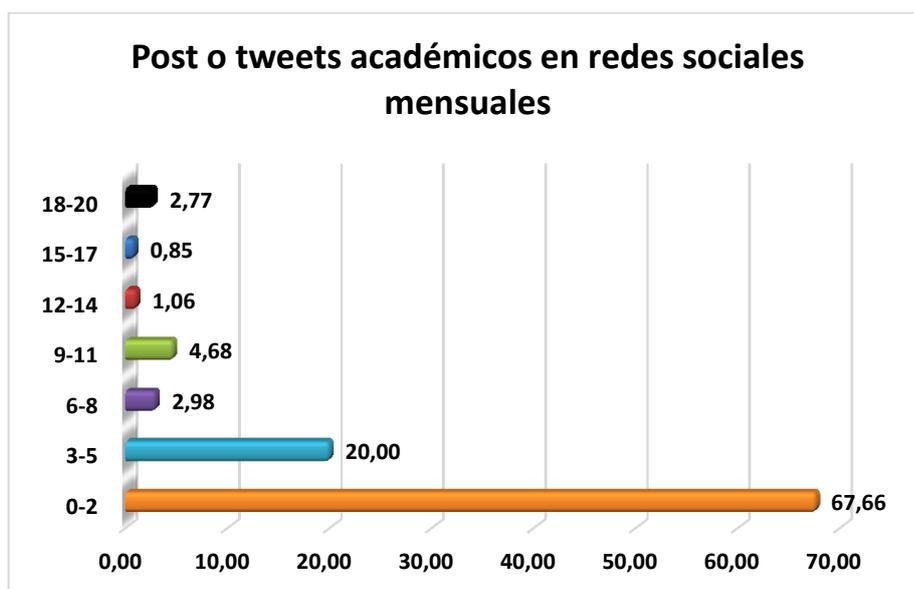


Figura 14: Post o Tweets académicos en redes sociales mensuales
Fuente: Elaboración propia

Como se puede visualizar en la Figura 15, los alumnos chatean un 39,36% entre 0 y 2 horas por temas académicos cada mes, seguidos de un 32,77% que chatean entre 3 y 5 horas mensuales. Es decir, la mayoría de alumnos chatean menos de 6 horas al mes por temas académicos.

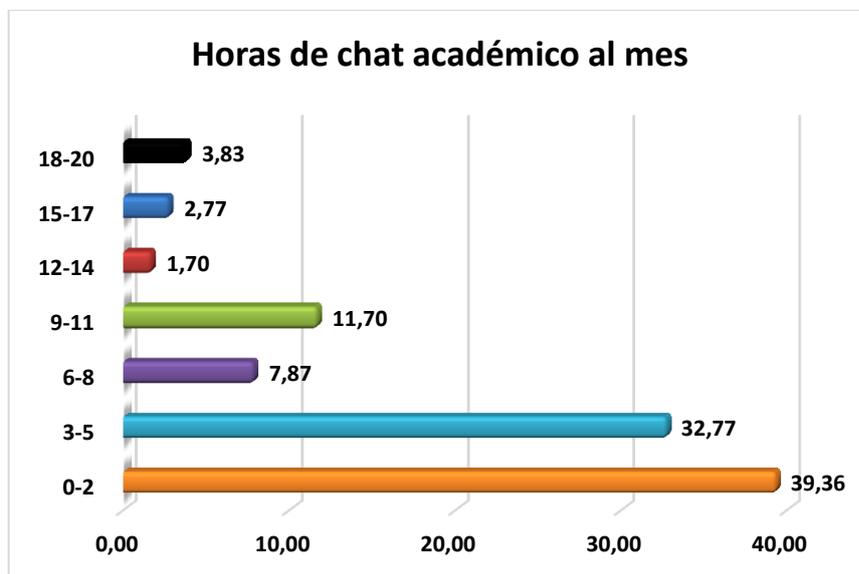


Figura 15: Horas de chat académico mensuales
Fuente: Elaboración propia

De los resultados obtenidos se determina que el 67,45% de los alumnos pasan hasta 14 horas mensuales indagando información de carácter académico en Internet; un 17,66% pasan de 15 a 29 horas en Internet investigando; además, se observa que la moda es 10 horas mensuales (18,5). Ver Figura 16.

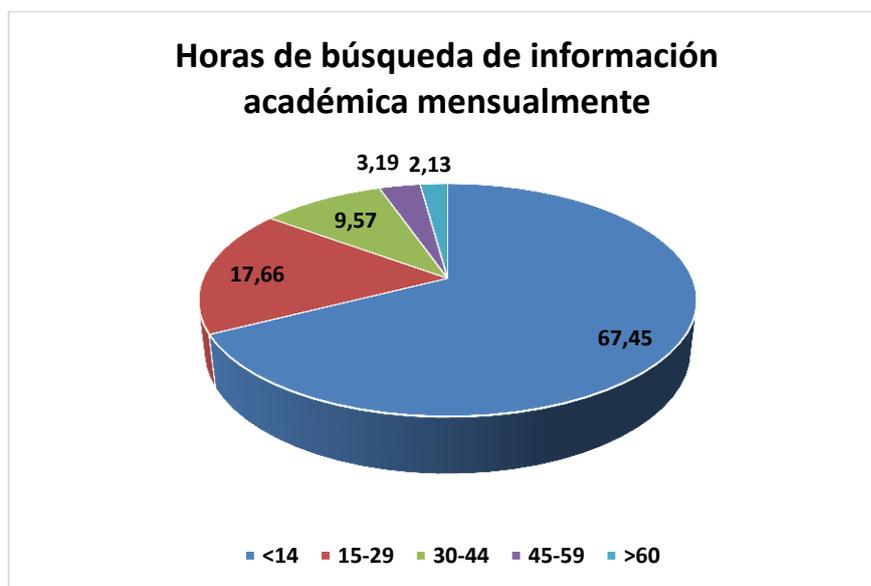


Figura 16: Horas de búsqueda de información académica en Internet por mes
Fuente: Elaboración propia

En la Figura 17, se puede observar que la biblioteca virtual de la Universidad Metropolitana es una herramienta poco utilizada por los alumnos, en vista de que el 36,4% de los alumnos no ingresan durante todo el mes.

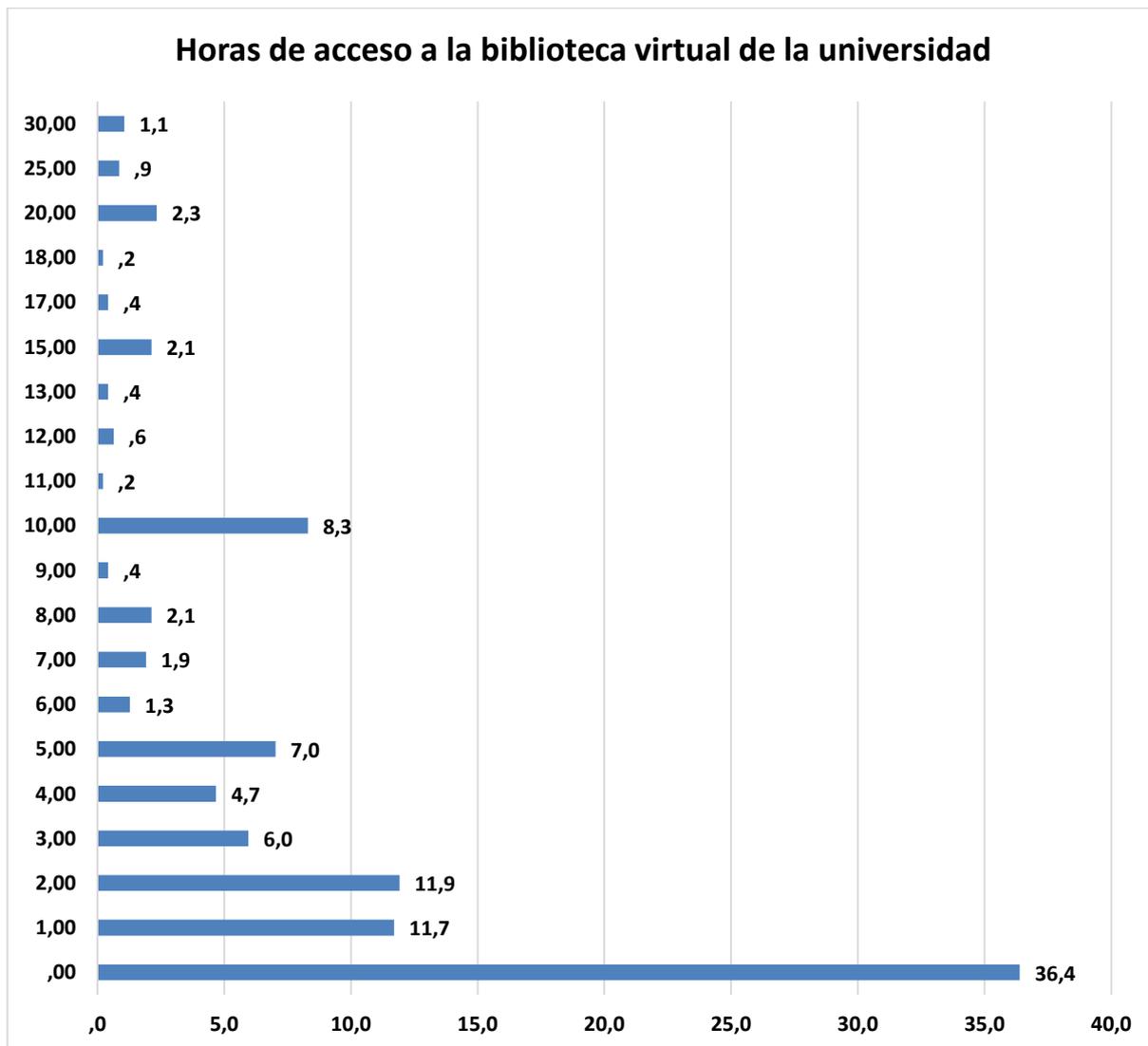


Figura 17: Horas de acceso a la biblioteca virtual de la universidad
Fuente: Elaboración propia

4.3.2. Usos de Internet en entretenimiento o diversión

A continuación se detallara los hallazgos realizados en lo referente a frecuencias de los alumnos encuestados para conocer los usos que le dan al Internet para acceder a información para entretenimiento y diversión.

En lo concerniente a los usos de Internet en el ámbito de entretenimiento, los alumnos de la Universidad Metropolitana de Quito utilizan un promedio de 10 horas semanales para

chatear por diversión; el 3,0% no utilizan el chat por temas de diversión. Como se puede ver en la Figura 18.

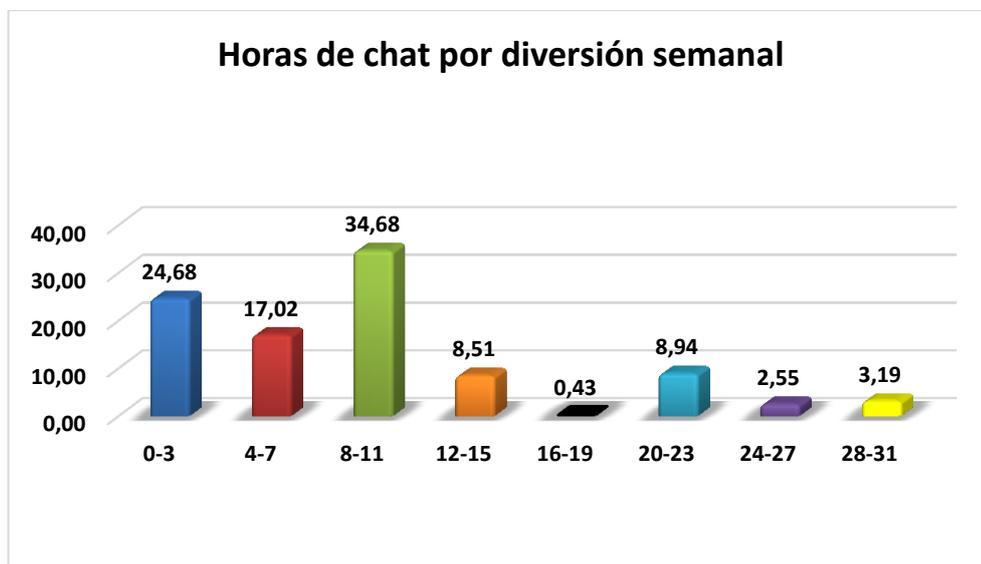


Figura 18: Horas de chat por diversión a la semana
Fuente: Elaboración propia

En lo referente a cuantas horas semanales los alumnos hacen uso de las redes sociales, se determinó que el 38,30% de los alumnos utiliza entre 7 y 13 horas a la semana redes sociales, un 2,1% de los alumnos no hace uso de redes sociales, además la moda es de 10 horas semanales (24,7%). Ver Figura 19.

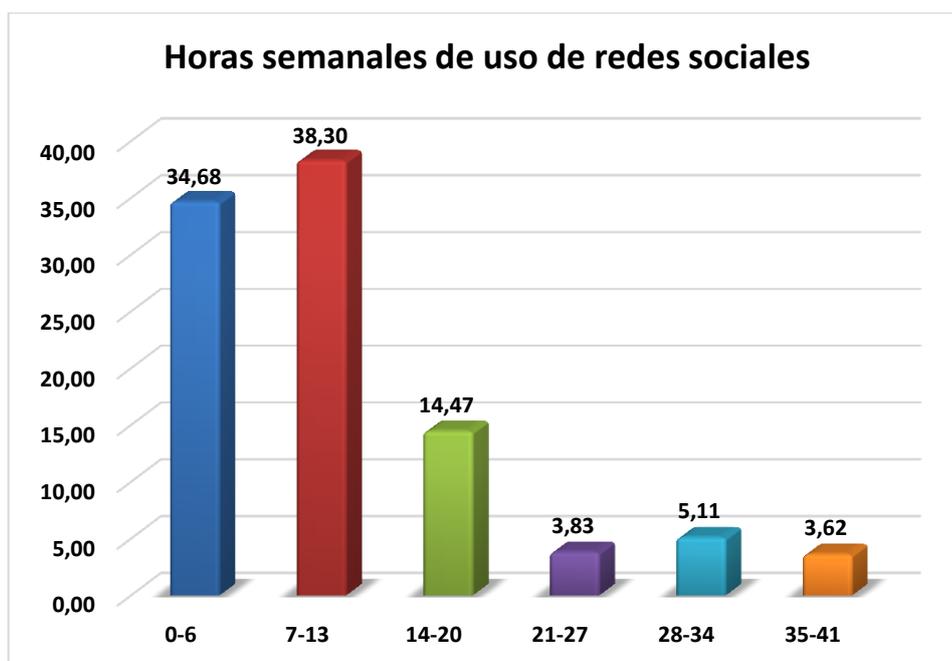


Figura 19: Horas semanales de uso de redes sociales
Fuente: Elaboración propia

De los alumnos encuestados el 47,2% no utilizan juegos en línea, el 12,1% usan juegos en línea más de 10 horas semanales. El promedio de uso de juegos en línea es de 3 horas.

El promedio de horas que usan los alumnos para la descarga de música, videos y programas es de 4 horas semanales, con una variación superior e inferior de 3,73. Ver Figura 20.

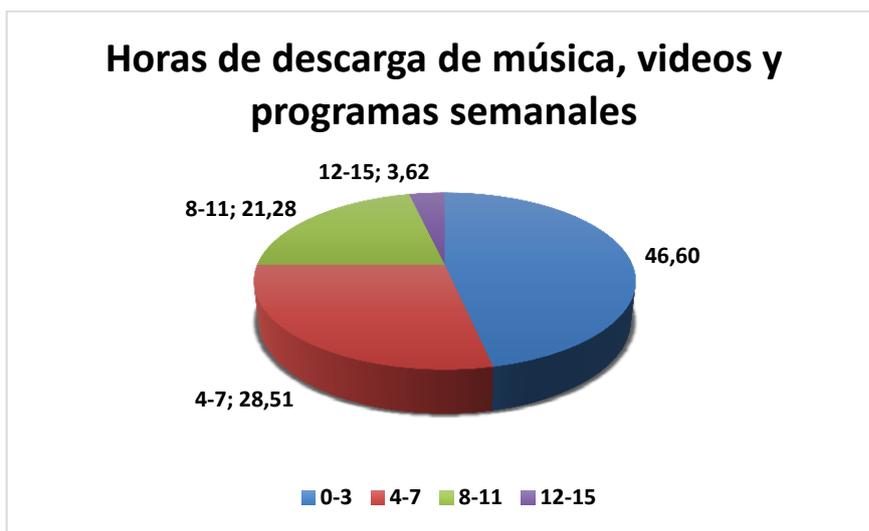


Figura 20: Horas semanales para descargar música, videos y programas
Fuente: Elaboración propia

Se determinó que los alumnos utilizan la herramienta YouTube para temas de diversión un promedio de 7 videos semanales, con una desviación estándar de 7,13. Como se puede ver en la Figura 21.



Figura 21: Videos de YouTube para diversión semanalmente
Fuente: Elaboración propia

4.4. Uso de redes sociales

En lo referente a los amigos en redes sociales (Twitter, Facebook y LinkedIn), se puede determinar que el 96,81% de los alumnos tienen menos de 500 seguidores en Twitter; el 69,79% tiene menos de 500 amigos en Facebook y 99,79% de los alumnos tienen menos de 500 contactos en LinkedIn. Como se puede ver en la Figura 22.

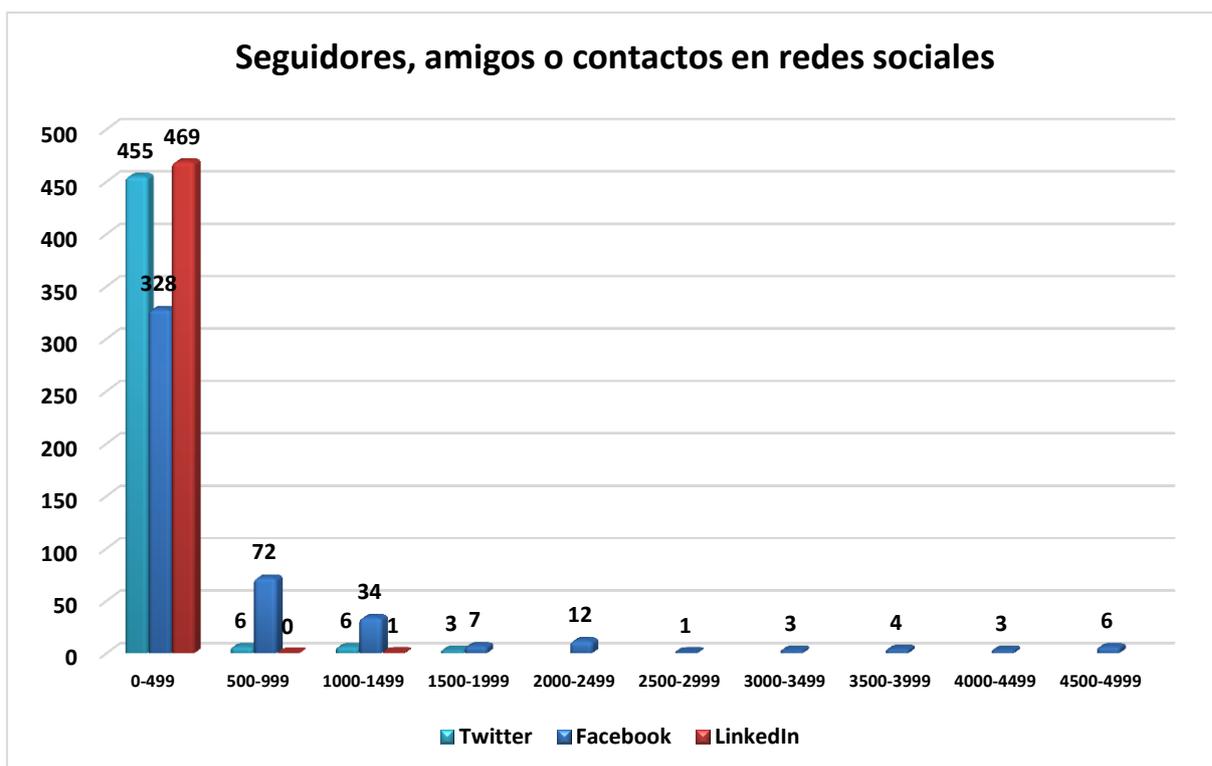


Figura 22: Seguidores, amigos o contactos en redes sociales
Fuente: *Elaboración propia*

4.5. Incidencias entre variables

Se ha determinado una relación significativa entre las variables edad y el lugar de conexión habitual a Internet (R de Pearson = 0,252, $p = 0,000$; $p < 0,05$) con un coeficiente de determinación del 6,35%. De igual manera se descubrió que existe relación entre la variable edad y el nivel de experiencia en el uso de Internet (R de Pearson = 0,372, $p = 0,000$; $p < 0,05$) con un 13,84% de coeficiente de determinación. Así mismo, presenta gran relación con el número de preguntas realizadas al docente mensualmente (R de Pearson = -0,184, $p = 0,000$; $p < 0,05$) presentando un coeficiente determinante del 3,39%.

La edad incide en cuantas horas los alumnos chatean por diversión (R de Pearson = -0,120, $p = 0,009$; $p < 0,05$) con un coeficiente de determinación del 1,44%; también se relaciona

con cuantas horas a la semana utiliza redes sociales (R de Pearson = -0,109, $p = 0,018$; $p < 0,05$) con 1,19% de coeficiente de determinación; además, presenta una relación significativa con cuántas horas descarga música, videos y programas (R de Pearson = -0,167, $p = 0,000$; $p < 0,05$) con un coeficiente de determinación del 2,79%.

De igual manera se encontró una relación significativa entre la edad y la cantidad de videos en YouTube para entretenimiento mensuales (R de Pearson = -0,178, $p = 0,000$; $p < 0,05$) con un coeficiente de determinación del 3,17%; la edad también se relaciona con cuantos amigos tiene en redes sociales (R de Pearson = -0,127, $p = 0,006$; $p < 0,05$) con un coeficiente de determinación del 1,61%.

Se determinó que el género se encuentra con una relación significativa con el nivel de conocimiento en el manejo de Internet (R de Pearson = -0,108, $p = 0,019$; $p < 0,05$) con un $R^2 = 1,67\%$; de igual manera se encuentra relacionado con el número de horas de conexión a Internet diarias (R de Pearson = 0,114, $p = 0,013$; $p < 0,05$) con un $R^2 = 1,30\%$; el género también se relaciona con cuantas horas a la semana accede a la biblioteca virtual de la universidad (R de Pearson = -0,112, $p = 0,015$; $p < 0,05$) con un $R^2 = 1,25\%$;

Se determina que el género se relaciona con el número de horas de usos de juegos en línea (R de Pearson = -0,132, $p = 0,004$; $p < 0,05$) con un $R^2 = 1,74\%$, determinando que los que más utilizan juegos en líneas son los hombres.

Los niveles de ingreso de las familias de los alumnos se encuentra relacionado con el lugar de conexión (R de Pearson = 0,154, $p = 0,001$; $p < 0,05$) con un $R^2 = 2,37\%$. Así mismo, presenta una relación con el número de días de conexión por semana (R de Pearson = -0,240, $p = 0,000$; $p < 0,05$) con un coeficiente de determinación del 5,76%; también los ingresos familiares están relacionados con el nivel de conocimiento en el manejo de Internet (R de Pearson = 0,112, $p = 0,015$; $p < 0,05$) con un $R^2 = 1,25\%$ y con la cantidad de horas de conexión al día (R de Pearson = 0,141, $p = 0,002$; $p < 0,05$) con un coeficiente de determinación del 1,99%.

4.6. Factorización: reducción de variables

Al aplicar análisis factorial a las variables del uso de Internet en actividades académicas y el uso de Internet en actividades de entretenimiento, se realizó pruebas con todas las variables, de igual manera se fue intercalando el uso de variables para determinar si es posible aceptar el análisis factorial. Por lo cual, se determinó que en ninguno de los casos se puede dar por cumplido el análisis factorial para las variables de índole académico. Ver Anexo 2, 3 y 4.

De la misma manera se realizó análisis factorial para las variables de entretenimiento, combinando todas sus variables e intercalando las mismas, así mismo, se determina que no se puede dar por cumplido el análisis factorial. Ver Anexo 5, 6 y 7.

Como se pudo determinar en los análisis factoriales de las variables tanto académicas como de entretenimiento no se puede dar por aceptado el análisis factorial en vista que el porcentaje de explicación de cada uno de los componentes no presenta un nivel aceptable.

4.7. Clusterización: determinación de perfiles

“El término «análisis de clúster» (o conglomerados) define una gran diversidad de técnicas y métodos cuyo fin es lograr una clasificación o agrupamiento de n individuos en g grupos según el comportamiento de éstos en una serie de variables” (Díaz de Rada, 1998).

4.7.1. Perfil académico

Para la generación de grupos para determinar los perfiles académicos, se utilizó k-medias, en este procedimiento se aplicaron para 2, 3, 4 y 5 agrupaciones, utilizando la técnica jerárquico ascendente (Ver Anexos 8, 9, 10 y 11), como variables de origen para el análisis clúster se utilizó las variables académicas que presentan mayor mediana (Tabla 7).

Tabla 7: Medianas de las variables académicas

Variable	Mediana
¿Cuántas veces a la semana ingresa a la plataforma virtual de su universidad?	2,00
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	4,00
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	4,00
¿Aproximadamente cuántos recursos educativos descarga de la plataforma virtual cada mes?	2,00
¿Aproximadamente cuántos videos académicos mira en YouTube cada mes?	3,00
¿Aproximadamente en cuántos foros virtuales participa cada mes?	1,00
¿Aproximadamente cuántos post o tweets sobre temas académicos realiza en las redes sociales por mes?	1,00
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	3,00
¿Aproximadamente cuántas horas busca información académica en Internet cada mes?	10,00
¿Aproximadamente cuántas horas utiliza la biblioteca virtual de la universidad cada mes?	2,00

Fuente: Elaboración propia

Utilizando análisis discriminante como explica Díaz de Rada (1998), se obtuvo la siguiente tabla de exactitud para cada grupo. Ver Anexo 12, 13, 14 y 15.

Tabla 8: Exactitud de clasificación de grupos académicos

Grupos	Exactitud
2	98,7 %
3	98,3 %
4	97,9 %
5	96,4 %

Fuente: Elaboración propia

Con los datos de la tabla anterior, luego de haber determinado el análisis discriminante para todos los tipos de grupos de clasificación, se determina que el grado de mayor exactitud es 98,7% que corresponde a la generación de 2 grupos. Además de obtener el nivel de exactitud más elevado, la clasificación en dos grupos ofrece una lectura más sencilla y fácil de explicar.

4.7.2. Perfil de entretenimiento

Para la creación de la variable el uso de Internet en actividades entretenimiento, se hizo uso de las variables originales de la pregunta N° 11 (Ver Anexo 1), que se refieren a las horas de: chat por diversión, uso de redes sociales y uso de juegos en línea; así como horas de descarga de música y cuántos videos para entretenimiento ve en YouTube.

Como recomienda Díaz de Rada (1998), para la generación de grupos se utilizó k-medias, para agrupaciones de 2, 3, 4 y 5 grupos. Ver Anexo 18, 19, 20 y 21.

Para determinar los conglomerados, se utilizó como variables de origen las variables de la pregunta N° 11 del anexo 1, de igual manera se realizó análisis discriminante para determinar el nivel de exactitud de clasificación para cada grupo (Ver Anexo 22, 23, 24 y 25), dando como niveles de exactitud los siguientes.

Tabla 9: Porcentaje de exactitud de grupos de uso de Internet para entretenimiento

Grupos	Exactitud
2	98,9 %
3	98,5 %
4	96,6 %
5	99,1 %

Fuente: *Elaboración propia*

En la Tabla 9, se puede ver que el mayor nivel de exactitud en la clasificación de grupos está dada para el grupo 5 con un nivel de exactitud del 99,1%; pero al momento de realizar un análisis gráfico y descriptivo de los grupos, la clasificación en 5 grupos se torna confusa y difícil de explicar; por tal motivo se ha seleccionado la clasificación en 2 grupos con un nivel de exactitud del 98,9%.

4.8. Descripción de perfiles

4.8.1. Académico

Con los valores de los procesos antes realizados, se determina que el grupo con mayor exactitud es el clasificado en dos grupos, cada grupo está determinado por: el número de preguntas al docente, número de preguntas a sus compañeros, cantidad de videos académicos en YouTube, horas de chat sobre temas académicos y horas de búsqueda de información académica.

De tal manera, que el uso de Internet para actividades académicas, queda clasificado en dos grupos, los cuáles se puede observar en la Figura 23.

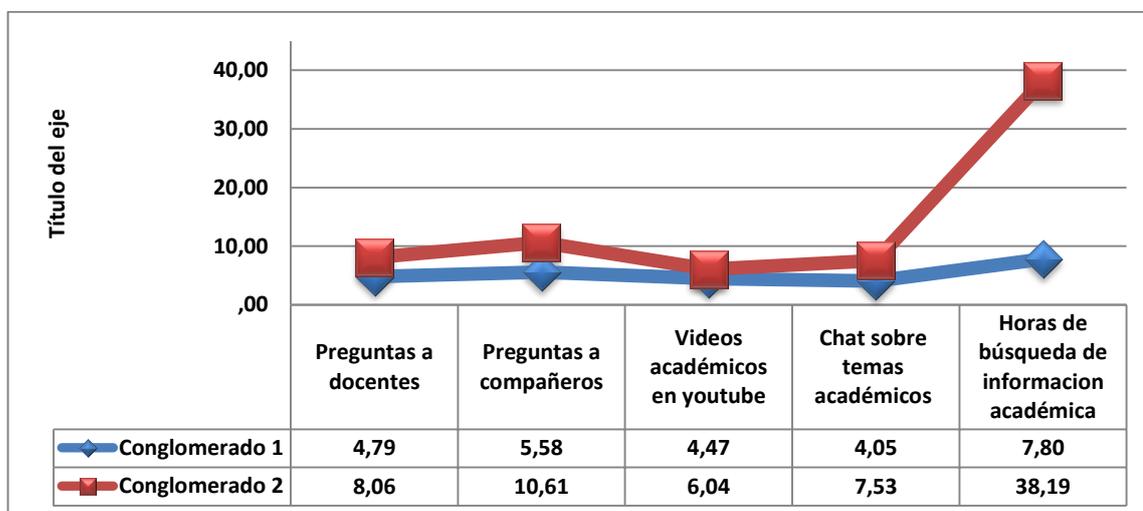


Figura 23: Centroides de los grupos del uso de Internet en actividades académicas
Fuente: Elaboración propia

El primer grupo está conformado por el 17,66% de los estudiantes, y presentan medias más bajas que el grupo dos en todos sus factores, a este grupo se lo denomina **Despreocupado**, porque realizan menos consultas tanto a docentes como a compañeros; miran menos videos académicos en YouTube; chatean menos horas sobre temas académicos y buscan menos tiempo información académica en la web. Este grupo está conformado por el 47,5% de mujeres y el 52,5% de hombres (Ver Anexo 16). La edad media de este grupo es de 23,33 años de edad (Ver Anexo 17). Los ingresos del núcleo familiar de este grupo tienden a ser de hasta 1000,00 dólares mensuales. El lugar habitual de conexión del grupo Despreocupado es desde su casa.

Mientras tanto el grupo dos, está conformado por el 82,34%, que se caracteriza por poseer valores más altos en todos los factores con referencia al primer grupo, se lo denomina **Preocupado**, porque tiende a realizar más consultas a docentes y compañeros; observa más videos en YouTube sobre temas académicos; chatea más horas sobre temas académicos y busca más tiempo información académica en la web. El 44,6% de este grupo son mujeres y el 55,4% son varones (Ver Anexo 16). La edad media de este grupo es de 23,45 años (Ver Anexo 17). Sus ingresos familiares tienden a ser iguales que el primer grupo, con ingresos de hasta 1000 dólares. Al igual que el grupo despreocupado el grupo preocupado tiene como lugar de conexión habitual su hogar.

La carrera con mayor cantidad de alumnos en el grupo Preocupado es: la carrera de derecho con el 24,1%, seguida de las carreras de ingeniería en sistemas de información y optometría con un 20,5% cada una; gestión empresarial con 14,5%; gestión de empresas turísticas y hoteleras 10,8%; ciencias administrativas y contables 4,8%; diseño gráfico 3,6% y por último talento infantil 1,2%.

En el nivel de conocimiento de uso de Internet, el nivel Despreocupado presenta una media de 7,66/10 mientras que el grupo Preocupado 7,82/10, se puede observar que la diferencia es leve entre los dos grupos.

4.8.2. Entretenimiento

La clasificación en dos grupos para la variable usos de Internet para entretenimiento está dado por: horas de chat por diversión semanales, horas semanales de utilización de redes sociales, hora a la semana de utilización de juegos en línea, las horas a la semana dedicadas a descarga música, videos y programas, y cuántos videos para entretenimiento mira en YouTube por semana. Con estas variables se logró determinar clasificaciones para dos grupos cómo se puede observar en la Figura 24.

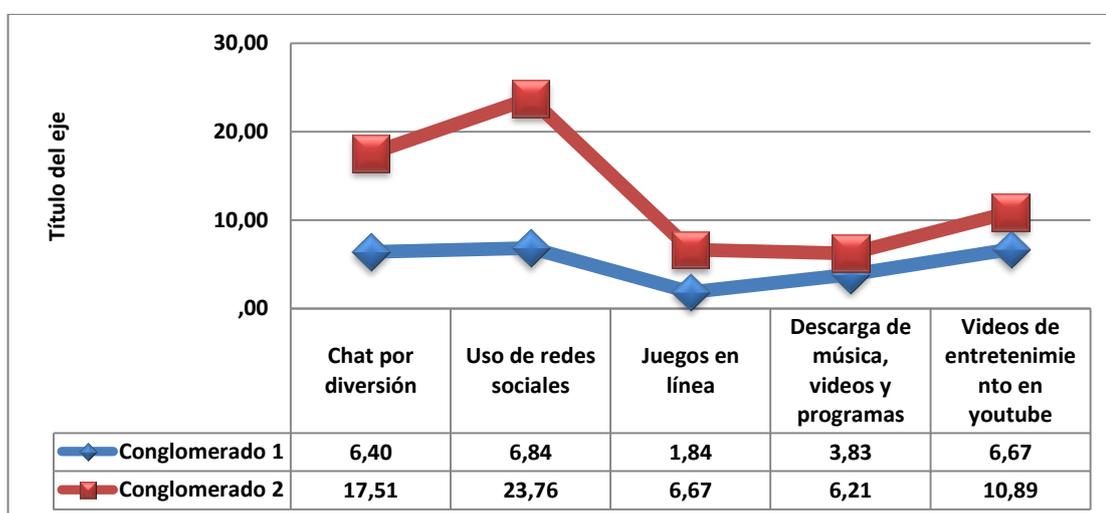


Figura 24: Centroides de los grupos del uso de Internet para entretenimiento
Fuente: Elaboración propia

Los usos de Internet para entretenimiento está clasificado en dos grupos, el primer grupo está conformado por el 26.17% de la población estudiantil, y como se puede ver en la Figura 24, presenta medias más bajas en relación al segundo grupo, por ende se ha visto conveniente denominarlo **Pasivo**, este grupo corresponde a los alumnos que menos horas chatean por diversión, hacen menor uso de redes sociales, menor uso de juegos en línea,

realiza menos descargas de música, videos y programas, además, visualizan menos cantidad de videos en YouTube para entretenimiento o diversión. Este grupo está conformado por el 52,8% de hombres y un 47,2% de mujeres (Ver Anexo 26). Presenta una edad media de 22 años de edad (Ver Anexo 27). Los ingresos mensuales de las familias de los alumnos pertenecientes a este grupo son hasta 1500 dólares mensuales. El grupo pasivo regularmente se conecta a Internet desde su casa.

El segundo grupo está conformado por el 73,83% de la población estudiantil, y presenta mayores índices de interacción en todos sus niveles como son: chat por diversión, mayor uso de redes sociales, mayor tiempo de utilizar juegos en línea, mayor cantidad de archivos de descargas en lo que concierne a música, videos y programas; y visualiza una mayor cantidad de videos en YouTube para entretenimiento o diversión, por eso se ha procedido a denominarlo **Activo**. El porcentaje de mujeres que conforman este grupo es del 47% y de hombres el 53% (Ver Anexo 26). La edad media para este grupo es de 23 años (Ver Anexo 27). Los ingresos familiares mensuales del grupo Activo superan los 1500 dólares. Al igual que el grupo pasivo tiende a conectarse habitualmente desde sus hogares.

4.9. Rendimiento académico

Para conocer el rendimiento académico de los alumnos de la Universidad Metropolitana de Quito, se operó las variables cuantas materias se matriculo el año anterior y cuantas materias aprobó, con esto se logró determinar el grado de éxito académico de los alumnos; en vista de que el periódico académico es anual, los alumnos pueden tomar entre 1 y 9 materias al año. Los alumnos tienden a matricularse entre 6 y 7 materias por periodo académico con un porcentaje del 50%.

Se determina que el 68,72% de los alumnos pertenecen al grupo de alto rendimiento el cual ha aprobado el 100% de las materias matriculadas, y se lo denominará grupo **Destacado**; mientras que el 31,28% corresponde a los alumnos que han reprobado al menos una materia en el año anterior, el cual se lo denominará grupo **Común** como lo muestra la Tabla 10.

Tabla 10: Grupos de rendimiento académico

Grupo	Frecuencia	Porcentaje
Destacado	323	68,72
Común	147	31,28
Total	470	100,00

Fuente: Elaboración propia

4.10. Comprobación de hipótesis

Luego de determinar los grupos o perfiles de uso de Internet para fines académicos; los usos de Internet para entretenimiento y diversión; y el nivel de rendimiento académico de los alumnos de la UMET, se procede a realizar la comprobación de las hipótesis del proyecto aplicando regresión logística.

4.10.1. El nivel de ingresos determina como se utiliza Internet para el aprendizaje

Para la comprobación de la hipótesis 1 se aplicó regresión logística binomial en vista de que la variable dependiente a tratar es de tipo dicotómica. Como variable independiente se tiene los ingresos de las familias de los estudiantes categorizados en 5 niveles y como variable dependiente los usos de Internet en actividades académicas: despreocupado y preocupado. Ver Anexo 28.

Aplicando regresión logística binomial se muestra la varianza que explica el modelo representado por el valor de Nagelkerke = 0,006 (0,6%). La prueba Wald presenta que ninguno de los niveles de ingresos tiene significancia sobre los grupos despreocupado y preocupado ($p > 0,05$) (Tabla 11).

Tabla 11: Coeficiente de regresión logística para los ingresos y el uso de Internet en actividades académicas

		Variables en la ecuación						I.C. 95% para EXP(B)	
		B	E.T.	Wald	gl	Sig.	Exp(B)	Inferior	Superior
Paso 1 ^a	Ing			1,834	4	,766			
	ing(1)	,453	,399	1,285	1	,257	1,572	,719	3,440
	ing(2)	,163	,367	,198	1	,656	1,177	,574	2,416
	ing(3)	,054	,360	,023	1	,881	1,056	,521	2,138
	ing(4)	-,049	,388	,016	1	,899	,952	,445	2,038
	Constante	-1,642	,258	40,672	1	,000	,194		

a. Variable(s) introducida(s) en el paso 1: ing.

Fuente: Elaboración propia

La prueba de ajuste de Hosmer y Lemeshow presenta ($\chi^2 = 0$, $p = 1$) en vista de que el $p > 0,05$, se determina que el modelo es ajustable a los datos; es decir los valores observados son idénticos a los predichos.

La prueba de Omnibus presenta ($\chi^2 = 1,769$, $p = 0,778$; $p > 0,05$) por lo que se determina que el modelo no es aceptable.

4.10.2. El nivel de ingresos determina como se utiliza Internet para entretenimiento

La hipótesis 2 se comprueba aplicando regresión Logística binomial en vista de que la variable dependiente a tratar es de tipo dicotómica. Como variable independiente se tiene los ingresos de las familias de los estudiantes categorizados en 5 niveles y como variable dependiente los usos de Internet en actividades de entretenimiento o diversión: activo o pasivo (Ver Anexo 29).

Aplicando regresión logística binomial se muestra la varianza que explica el modelo representado por el valor de Nagelkerke = 0,035 (3,5%). La prueba Wald presenta significancia en el ingreso 4 (hasta 1500 dólares) sobre los grupos activo y pasivo ($p = 0,039$; $p < 0,05$) (Tabla 12).

Tabla 12: Coeficiente del modelo de regresión logística para los ingresos y los usos de Internet en entretenimiento

		Variables en la ecuación						I.C. 95% para EXP(B)	
		B	E.T.	Wald	gl	Sig.	Exp(B)	Inferior	Superior
Paso 1 ^a	ing			11,237	4	,024			
	ing(1)	-,103	,375	,075	1	,784	,902	,432	1,883
	ing(2)	-,462	,345	1,794	1	,180	,630	,320	1,239
	ing(3)	,035	,307	,013	1	,909	1,036	,567	1,891
	ing(4)	,635	,307	4,262	1	,039	1,886	1,033	3,446
	Constante	-1,087	,219	24,722	1	,000	,337		

a. Variable(s) introducida(s) en el paso 1: ing.

Fuente: Elaboración propia

La prueba de ajuste de Hosmer y Lemeshow presenta ($\chi^2 = 0$, $p = 1$) en vista de que el $p > 0,05$, se determina que el modelo es ajustable a los datos; es decir los valores observados son idénticos a los predichos.

La prueba de Omnibus presenta ($\chi^2 = 11,324$, $p = 0,023$; $p < 0,05$) por lo que se determina que el modelo es aceptable para los presentes datos.

4.10.3. El uso de la tecnología en el aprendizaje incide en el rendimiento académico

Para la comprobación de la hipótesis 3 primeramente se determinó la variable dependiente rendimiento académico (destacado, común).

Aplicando regresión logística binomial se presenta la explicación del modelo Nagelkerke = 0,014 (1,4%). La prueba Wald presenta significancia sobre el rendimiento académico ($p = 0,040$; $p < 0,05$) (Tabla 13).

Tabla 13: Coeficiente de regresión logística entre usos de Internet para actividades académicas y rendimiento académico

		Variables en la ecuación							
		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
Paso 1 ^a	perfil_academico(1)	,591	,287	4,228	1	,040	1,806	1,028	3,171
	Constante	-1,284	,266	23,241	1	,000	,277		

a. Variable(s) introducida(s) en el paso 1: perfil_academico.

Fuente: Elaboración propia

La prueba de ajuste de Hosmer y Lemeshow presenta ($\chi^2 = 0$, $p = 1$) en vista de que el $p > 0,05$, se determina que el modelo es ajustable a los datos; es decir los valores observados son idénticos a los predichos.

La prueba de Omnibus presenta ($\chi^2 = 4,552$, $p = 0,033$; $p < 0,05$) por lo que se determina que el modelo se debe de aceptar.

Se determina que el uso de Internet en actividades académicas incide en el rendimiento académico de los alumnos de la UMET.

4.10.4. El uso de la tecnología para entretenimiento incide en el rendimiento académico.

Como variable dependiente para la comprobación de la hipótesis 4 se hizo uso de la variable rendimiento académico (destacado, común), como es una variable dicotómica se aplicó regresión logística binaria.

Presenta un explicación del Modelo Nagelkerke = 0,020 (2%). Se puede observar que la prueba de Wald presenta significancia relevante de la variable independiente (usos de Internet en entretenimiento) sobre la dependiente (rendimiento académico) ($p = 0,009$; $p < 0,05$) (Tabla 14).

Tabla 14: Coeficiente de regresión logística para usos de Internet en entretenimiento y rendimiento académico

		Variables en la ecuación						I.C. 95% para EXP(B)	
		B	E.T.	Wald	gl	Sig.	Exp(B)	Inferior	Superior
		Paso 1 ^a	perfil_entretenimiento(1)	-,568	,219	6,728	1	,009	,566
	Constante	-,378	,184	4,250	1	,039	,685		

a. Variable(s) introducida(s) en el paso 1: perfil_entretenimiento.

Fuente: Elaboración propia

Al aplicar la prueba de ajuste de Hosmer y Lemeshow refleja ($\chi^2 = 0$, $p = 1$) en vista de que el $p > 0,05$, se determina que el modelo es ajustable a los datos; ya que el grado de datos predichos concuerda con los datos observados.

La prueba de Omnibus presenta ($\chi^2 = 6,623$, $p = 0,010$; $p < 0,05$) por tanto se determina que el modelo es predecible a los datos.

En relación a los datos anteriores se determina que el uso de Internet en actividades de entretenimiento o diversión incide sobre el rendimiento académico de los estudiantes.

CAPÍTULO V

ANÁLISIS Y DISCUSIÓN DE RESULTADOS

En el presente capítulo, se analiza los datos observados y se procede a contrastar con las diferentes preguntas e hipótesis del trabajo de investigación.

5.1. Análisis hipótesis 1

Al analizar los datos observados de la hipótesis 1 se determina que los ingresos económicos no determinan los usos de Internet en actividades académicas, en vista de que no presentan ninguna relación significativa de la variable independiente (ingresos económicos) sobre los la variable dependientes (usos de Internet en actividades académicas).

Para conocer la diferencia entre los valores predichos y los valores observados, presenta ($\chi^2 = 1,769$, $p = 0,778$; $p > 0,05$) con 4 grados de libertad; los valores arrojados presentan que el nivel de ingresos NO determina como se utiliza Internet para entretenimiento.

Aunque en el presente proyecto no se logra dar por cumplida la hipótesis, los autores Torres & Infante (2011), en cambio determinan que mientras mayores sean los ingresos familiares mayor será la probabilidad de acceder a un computador y tener acceso a Internet.

Los ingresos familiares de los alumnos de la UMET se encuentran con mayor concentración entre 1000 y más de 1500 dólares mensuales; y con 65,1% de conexión a Internet desde sus hogares. Se deduce que aunque los alumnos cuentan con altos ingresos económicos mensuales y que casi 2/3 de la población estudiantil poseen computador y acceso a Internet, y se conectan por lo general 7 días a la semana, los usos que le dan al Internet no van encaminado a actividades académicas.

5.2. Análisis hipótesis 2

Para determinar si el nivel de ingresos determina como se utiliza Internet para entretenimiento, se observa que a prueba de Wald arroja una relación de significancia en lo referente al ingreso 4 (más de 1500 dólares) y el uso de Internet en actividades de entretenimiento, ($p = 0,039$; $p < 0,05$); pero en los demás coeficientes no presentan coeficientes significantes de relación ($p > 0,05$) y con una representación del modelo de 3,5%.

A pesar de presentar un coeficiente con significancia de la variable independiente sobre la dependiente, se procede a determinar que el nivel de ingresos NO determina como se utiliza Internet para entretenimiento.

De los datos observados se determina, que el nivel socioeconómico de los alumnos de la UMET se encuentra entre el nivel medio y alto, en vista de los ingresos de su núcleo familiar. Así mismo, se determina que aunque en el nivel más alto de ingresos económicos

(más de 1500 dólares mensuales) tiene una relación significativa sobre el uso de Internet en actividades de entretenimiento, no se puede dar por cumplida la hipótesis ya que los demás coeficientes no muestran significancia alguna de la variable independiente sobre la dependiente ($p < 0,05$).

El modelo para los datos es aceptable en vista de que los datos observados son idénticos a los valores predichos ($\chi^2 = 11,324$, $p = 0,023$; $p < 0,05$).

5.3. Análisis hipótesis 3

El uso de Internet en el aprendizaje y la incidencia que presenta sobre el rendimiento académico, muestra relación significativa.

La prueba de chi-cuadrado arroja ($\chi^2 = 4,552$, $p = 0,033$; $p < 0,05$), por lo tanto el modelo es deducible para los datos, y se visualiza que existe una incidencia significativa de la variable independiente sobre la dependiente.

El uso de Internet en actividades de aprendizaje SI determina el nivel de rendimiento académico de los alumnos, por lo tanto se deduce que mientras más tiempo y recursos invierte o dedica el alumno en actividades de aprendizaje, mayor será su nivel de rendimiento académico.

Al aprobar la incidencia del uso del Internet en actividades académicas sobre el rendimiento del alumno se concuerda con Torres (2012), el cual sostiene que los usos del Internet presentan efectos positivos para el rendimiento académico, pero se difiere con Ayala (2007) el cual determina que estos dos tipos de variables no muestran incidencia la una sobre la otra.

También se difiere con Castaño (2011), el cual menciona que el hecho de utilizar Internet para actividades académicas no se transforma en un mejor rendimiento académico.

Por lo tanto, se determina que los alumnos del grupo de los preocupados tienden a tener una probabilidad más alta de presentar mayor éxito académico, y los alumnos despreocupados tienden a perder una o más materias en un periodo académico.

Con lo antes expuesto se determina que los alumnos con mejor éxito académico (destacados) son los que mayor tiempo invierten en consultas a los docentes, a sus compañeros, observan más videos académicos en la herramienta YouTube, chatean más

sobre temas académicos y pasan un mayor tiempo buscando información de carácter académico. A diferencia de los alumnos con menor rendimiento académico (común) los cuales en todos los aspectos presentan menor interacción en comparación a los alumnos destacados.

5.4. Análisis hipótesis 4

En cuanto al uso de Internet para entretenimiento y su incidencia en el rendimiento académico se observa:

Que el modelo se acepta en vista de que los datos observados son iguales a los esperados ($\chi^2 = 0$, $p = 1$) y el modelo es predecible para los datos.

En base a los resultados del capítulo IV se determina que el uso de Internet para entretenimiento SI incide en el rendimiento académico. Por ende se determina que los alumnos activos tienden a presentar un bajo rendimiento académico, a diferencia de los alumnos pasivos que presentan mejor éxito académico.

Los alumnos con bajo rendimiento académico (común) son los alumnos que mayor tiempo pasan chateando, hacen uso de redes sociales, juegan en línea, descargan de música, videos y programas y visualizan mayor número de videos para entretenimiento en YouTube, a diferencia de los alumnos que presentan mejor rendimiento académico (destacados) los cuales hacen menos usos de Internet en actividades para entretenimiento.

Los alumnos de la UMET presentan ingresos económicos de Hasta 1000 (23,8%) y más de 1500 dólares (23,6%), y se conectan regularmente desde sus hogares, 7 días a la semana y con una conexión media de 4 horas diarias.

Se determinó que los la edad es un determinante muy importante en relación a la experiencia en Internet, nivel de conocimiento, tiempo de acceso a Internet y la forma de utilizar Internet tanto en actividades académicas como de entretenimiento.

Los ingresos económicos no determinan los usos de Internet en actividades académicas, puesto que los datos no se ajustaron la modelo de regresión logística binomial, de igual manera, que los ingresos económicos no determinan los usos de Internet para entretenimiento.

Los usos de Internet para actividades académicas, SI incide sobre el nivel de rendimiento académico de los alumnos, ya que los datos se ajustan al modelo de regresión logística. Se

determinó que mientras más tiempo invertido en el uso de Internet para actividades académicas, mayor nivel de éxito académico logran los alumnos.

Los usos de Internet para entretenimiento, SI presentan incidencia sobre el nivel académico de los estudiantes. El modelo de regresión logística binomial se ajusta a los datos. Y se determina que mientras más tiempo se dedica a usos de Internet para entretenimiento mayores serán las probabilidades de presentar menor rendimiento académico.

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

Por el levantamiento de la información se puede observar que:

- El uso de la tecnología para actividades académicas como para entretenimiento inciden en el rendimiento académico de los estudiantes de la Universidad Metropolitana, en vista que el modelo de regresión logística binomial determina que al realizar el cálculo de la variable independiente sobre la dependiente esta tiende a cambiar.
- En el Ecuador en cuanto al acceso a Internet tiende a reducirse, en vista de que en la mayoría de los hogares de los alumnos cuentan con acceso a Internet, sin importar el nivel de ingresos económicos.
- El instrumento aplicado, está constituido con las variables apropiadas que permitieron llevar a cabo los procesos de minería de datos y comprobar las hipótesis planteadas.
- La minería de datos, permite extraer conocimiento a partir de los datos. La calidad de los datos observados, depende de la aplicación correcta de los modelos y determinación exitosa de cada uno de las fases de minería de datos.
- Los procesos de la fase de minería de datos, pueden ser ejecutados usando la Herramientas desarrollada por IBM SPSS, en vista de que presenta una interfaz amigable y predecible, es muy útil para trabajar con datos de muestras pequeñas o medianas.

Recomendaciones

- Desarrollar proyectos de investigación para determinar el nivel de incidencia de los usos de la tecnología tanto para actividades de aprendizaje y de entretenimiento y el nivel de excelencia académica en alumnos de instituciones de nivel básico y medio.
- Concienciar a los alumnos sobre las potencialidades académicas que se pueden dar con el uso de Internet, puesto su uso podría ayudar a lograr un mejor rendimiento académico. Aprovechando los recursos y herramientas ofertadas en la Web 2.0.
- Aplicar mayor tiempo y esfuerzo en las fases de minería de datos como son: recopilación, selección, limpieza y transformación de variables, para determinar datos de calidad, y de esta manera asegurar la calidad del conocimiento extraído
- Cuando se tenga grandes cantidades de datos es conveniente realizar la adquisición de software licenciado, para no tener problemas de inconsistencia de datos y el tratamiento adecuado de la información. Pues que, SPSS no es recomendable para grandes cantidades de datos aunque incorpora la mayoría de técnicas y herramientas de minería de datos.
- Para trabajo futuros sería conveniente reestructura la encuesta, para poder determinar nuevos índices o coeficiente de estudio.

BIBLIOGRAFÍA

- Andes, D. L. (2009). Una aplicación del método jerárquico de mezclas para la clasificación de los municipios venezolanos según variables socioeconómicas, (2).
- Ayala, C. (2007). *Relación entre el uso de Internet y el logro académico*.
- Boneu, J. M. (2007). Plataformas abiertas de e-learning para el soporte de contenidos educativos abiertos. *Learning Technology*, 4, 36–47. Obtenido de <http://www.uoc.edu/rusc/4/1/dt/esp/boneu.pdf>
- Bustos Sánchez, A., y Coll Salvador, C. (2010). Los entornos virtuales como espacios de enseñanza y aprendizaje. *Revista Mexicana de Investigación Educativa*, 15, 163–184. Obtenido de <http://scielo.unam.mx/pdf/rmie/v15n44/v15n44a9.pdf>
- Castro, S. M., Clarenc, C. A., Lenz, C. L. de, Moreno, M. E., y Tosco, N. B. (2013). *Analizamos 19 Plataformas De E-Learning* (2013th ed., p. 154). Obtenido de <http://cooperacionib.org/191191138-Analizamos-19-plataformas-de-eLearning-primera-investigacion-academica-colaborativa-mundial.pdf>
- Castaño, J. (2011). El uso de internet para la interacción en el aprendizaje: un análisis para la eficacia y la igualdad en el sistema catalán. Obtenido de http://www.tesisred.net/bitstream/handle/10803/52561/Tesis_Jonatan_Casta%C3%B1o.pdf?sequence=1
- De la Fuente, F. (s.f.). Análisis De Conglomerados. *Fac. Ciencias Económicas Y Empresariales*. Obtenido de <http://www.fuenterrebollo.com/Economicas/ECONOMETRIA/SEGMENTACION/CONGLOMERADOS/conglomerados.pdf>
- Díaz de Rada, V. (1998). Diseño de tipologías de consumidores mediante la utilización conjunta del Análisis Cluster y otras técnicas multivariantes. *Economía Agraria*, 182, 75–104. Obtenido de http://www.magrama.gob.es/ministerio/pags/biblioteca/revistas/pdf_reea/r182_02.pdf
- DiMaggio, P., y Hargittai, E. (2001). From the “Digital Divide” to “Digital Inequality”: Studying Internet use as Penetration Increases. *Center for Arts and Cultural Policy Studies, Princeton University*, 15, 1–23. Obtenido de http://www.maximise-ict.co.uk/WP15_DiMaggioHargittai.pdf
- Druetta, D. D. M. C. (2006). ¿Es Internet Un Medio De Comunicación?, 7, 9.
- Gallardo, A. R. (2006). *La Brecha Digital y sus determinantes* (p. 254). Obtenido de http://132.248.242.3/~publica/archivos/libros/brecha_digital_y_determinantes.pdf
- Garbanzo Vargas, M. G. (2007). Factores asociados al rendimiento académico en estudiantes universitarios, una reflexión desde la calidad de la educación superior pública. *Revista Educación*, 31(1), 43–63. Obtenido de <http://www.revistas.ucr.ac.cr/index.php/educacion/article/view/1252>
- Garduño Vera, R. (2007). Caracterización del docente en la educación virtual: consideraciones para la Bibliotecología. *Investigación Bibliotecológica*, 21, 157–183.

- Gonzalez Pineda, J. A. (2003). El rendimiento escolar. Un analisis de las variables que lo condicionan. *Revista Galego-Portuguesa de Psicoloxía E Educación*, 8, 247–258. Obtenido de http://ruc.udc.es/bitstream/2183/6952/1/RGP_9-17.pdf
- Hargittai, E., y Shafer, S. (2006). Differences in actual and perceived online skills: The role of gender. *Social Science Quarterly*, 87(2), 432–448. doi:10.1111/j.1540-6237.2006.00389.x
- Hernández, J., Ramírez, J., y Ferri, C. (2007). *Introducción a la Minería de Datos*. Mexico: Ed. Pearson Educación.
- Herrera Castellanos, M. (2009). Formula para cálculo de la muestra poblaciones finitas. *Hospital de Roosevelt*. Obtenido de <https://investigacionpediahr.files.wordpress.com/2011/01/formula-para-cc3a1lculo-de-la-muestra-poblaciones-finitas-var-categorica.pdf>
- Ilabaca, J. S. (s.f.). Usos Educativos de Internet, 1–6. Obtenido de <http://users.dcc.uchile.cl/~jsanchez/Pages/papers/usoseducativosdelInternet.pdf>
- Inec. (2013). Tic 2013. *2013 8th Computing Colombian Conference (8CCC)*, 1–3. doi:10.1109/ColombianCC.2013.6637543
- Martínez H, E. (s.f.). Cluster: K-medias, 1–4. Obtenido de <http://www.uantof.cl/facultades/csbasicas/Matematicas/academicos/emartinez/magister/cluster1.pdf>
- Molina López, J. M., y Herrero, J. G. (2006). Técnicas de análisis de datos, 266. Obtenido de <http://www.giaa.inf.uc3m.es/docencia/II/ADatos/apuntesAD.pdf>
- Moreno, G. (2007). *Técnicas más usadas en Minería de Datos*. Obtenido de <https://gamoreno.wordpress.com/2007/10/03/tecnicas-mas-usadas-en-la-mineria-de-datos/>
- Nadu, T. (2011). Performance Analysis of Clustering Algorithms in Detecting Outliers, 2(1), 486–488. Obtenido de <http://www.ijcsit.com/docs/Volume2/vol2issue1/ijcsit2011020105.pdf>
- Pérez López, C., y Santin González, D. (2008). *MINERIA DE DATOS, TÉCNICAS Y HERRAMIENTAS* (p. 775). Madrid.
- Ramírez Ramírez, R., y Castillo Ramon, H. E. (2002). Guía para la aplicación de análisis multivariado a las encuestas de hogares. Obtenido de http://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib0512/Libro.pdf
- Sar, A. (2004). La Brecha del Conocimiento y la Brecha Digital, 1–7. Obtenido de http://www.littec.ungs.edu.ar/SSI2004/4_Sar.pdf
- Serrano, A., y Martínez, E. (2003). *La brecha digital. La Brecha Digital: Mitos y Realidades* (pp. 7–29). doi:10.1174/02109399960256739

- Torres, J. C. (2012). *Análisis de las relaciones entre los niveles de ingreso, edad y género de los estudiantes, los usos de Internet y el rendimiento académico en un grupo de universidades ecuatorianas presenciales*
- Torres, J. C., y Infante, A. (2011). *Desigualdad digital en la universidad, usos del Internet en el Ecuador.*
- Tusell, F. (2012). Análisis Multivariante. Obtenido de <http://www.et.bs.ehu.es/~etptupaf/nuevo/ficheros/estad4/multi.pdf>
- Villardón, J. L. V. (2011). El Análisis de Clusters, 22. Obtenido de <http://benjamindespensa.tripod.com/spss/AC.pdf>
- Voges, K. E. (2009). Cluster Analysis Using Rough Clustering and k-Means Clustering, (1991), 2009–2011. Obtenido de <http://www.irma-international.org/viewtitle/13629/>
- Weforum. (2014). *Global Information Technology Report 2014 - El Foro Económico Mundial. Weforum* (p. 369 p.). Obtenido de <http://reports.weforum.org/global-information-technology-report-2014/>

ANEXOS

Anexo 1

Encuesta realizada a estudiantes de la Universidad Metropolitana

Estimado estudiante, solicitamos su colaboración contestando esta encuesta, la que permitirá desarrollar una investigación para conocer el uso de Internet en las universidades del Ecuador.

1. Responda la siguiente pregunta										
¿En qué universidad estudia?										
2. Responda la siguiente pregunta										
¿Qué carrera estudia?										
3. Responda la siguiente pregunta										
¿Cuál es su edad?										
4. Responda la siguiente pregunta										
¿Cuál es su género?										
				Hombre			Mujer			
				()			()			
5. Los ingresos mensuales de su familia son de:										
Hasta 350 dólares										
Hasta 600 dólares										
Hasta 1.000 dólares										
Hasta 1.500 dólares										
Más de 1.500 dólares										
6. ¿Desde dónde se conecta habitualmente a Internet? (escoja solo una opción)										
Desde la casa										
Desde un cyber café										
Desde el trabajo										
Desde la Universidad										
Desde una red móvil (movistar, claro, cnt)										
7. Responda la siguiente pregunta										
De 1 a 7, ¿cuántos días a la semana se conecta Internet?										
1 2 3 4 5 6 7										
()										
8. Responda las siguientes preguntas										
De 1 a 10 su nivel de conocimientos en el manejo de Internet es:										
1 2 3 4 5 6 7 8 9 10										
()										
9. Responda las siguientes preguntas										
¿Aproximadamente cuántas horas se conecta cada día?										
()										
¿Hace cuántos años se conecta a Internet?										
()										
10. En lo referente a las asignaturas en las que está matriculado										
¿Cuántas veces a la semana ingresa a la plataforma virtual de su universidad?										
()										
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?										
()										
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?										
()										
¿Aproximadamente cuántos recursos educativos descarga de la plataforma virtual cada mes?										
()										
¿Aproximadamente cuántos videos académicos mira en youtube cada mes?										
()										
¿Aproximadamente en cuántos foros virtuales participa cada mes?										
()										
¿Aproximadamente cuántos post o tweets sobre temas académicos realiza en las redes sociales por mes?										
()										
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?										
()										
¿Aproximadamente cuántas horas busca información académica en Internet cada mes?										
()										
¿Aproximadamente cuántas horas utiliza la biblioteca virtual de la universidad cada mes?										
()										

11. En lo referente al entretenimiento y diversión en Internet	
¿Aproximadamente cuántas horas a la semana chatea por diversión?	()
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	()
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	()
¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	()
¿Aproximadamente cuántos videos para entretenimiento mira en YouTube cada semana?	()

12. Responda las siguientes preguntas	
¿Aproximadamente cuántos seguidores tiene en Twitter?	()
¿Cuántos amigos tiene en Facebook?	()
¿Cuántos contactos tiene en LinkedIn?	()

13. Responda con una X en SI o NO a las siguientes preguntas	SI	No
Tiene un blog	()	()
Tiene cuenta en YouTube	()	()
Tiene cuenta en www.del.icio.us	()	()

14. ¿Cuál es su nivel de uso de los siguientes dispositivos? (1 significa no usar y 10 significa utilizar al máximo)	1	2	3	4	5	6	7	8	9	10
Smartphone con cámara fotográfica y acceso a Internet	()	()	()	()	()	()	()	()	()	()
Teléfono móvil con acceso a Internet	()	()	()	()	()	()	()	()	()	()
Teléfono móvil sin acceso a Internet	()	()	()	()	()	()	()	()	()	()
Computador portátil	()	()	()	()	()	()	()	()	()	()
Tablet (iPad; Galaxy Tab, Kindle, etc.)	()	()	()	()	()	()	()	()	()	()
Cámara digital	()	()	()	()	()	()	()	()	()	()
iPod / MP3 Player	()	()	()	()	()	()	()	()	()	()

15. De 1 a 10 valore los siguientes aspectos (1 significa no estar de acuerdo y 10 estar completamente de acuerdo)	1	2	3	4	5	6	7	8	9	10
Internet le permite elaborar los trabajos más rápido y con menos esfuerzo	()	()	()	()	()	()	()	()	()	()
Usted confía en la información de Internet para realizar sus tareas	()	()	()	()	()	()	()	()	()	()
Internet le permite prescindir de la Biblioteca	()	()	()	()	()	()	()	()	()	()
Internet facilita el proceso de aprendizaje	()	()	()	()	()	()	()	()	()	()
Internet le permite mejorar sus calificaciones	()	()	()	()	()	()	()	()	()	()
Usted presenta trabajos académicos copiados desde Internet	()	()	()	()	()	()	()	()	()	()

16. Responda las siguientes preguntas referentes a sus profesores. (Se recomienda evaluar de forma general a todos sus profesores)	SI	NO	A veces
Su profesor ingresa a la plataforma virtual	()	()	()
Contesta sus consultas por correo electrónico	()	()	()
Chatea con usted eventualmente sobre aspectos académicos	()	()	()
Su profesor comenta en redes sociales sobre temas académicos	()	()	()
Le sube materiales digitales a la plataforma virtual	()	()	()
Le recomienda recursos digitales de la biblioteca virtual	()	()	()
Le recomienda videos sobre temas académicos	()	()	()
Le plantea cuestionarios o evaluaciones en la plataforma virtual	()	()	()
Le plantea foros virtuales	()	()	()
Su profesor tiene una página web, blog o perfil de Facebook	()	()	()
Su profesor tiene cuenta de Twitter	()	()	()

17. Responda las siguientes preguntas:	
En el semestre anterior, ¿en cuántas asignaturas se matriculó?	()
En el semestre anterior ¿cuántas asignaturas aprobó?	()

Anexo 2

Análisis Factorial de las variables académicas con 10 variables

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,801
Chi-cuadrado aproximado		697,573
Prueba de esfericidad de Bartlett	Gl	45
	Sig.	,000

Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3,010	30,104	30,104	3,010	30,104	30,104
2	1,214	12,138	42,241	1,214	12,138	42,241
3	,989	9,888	52,129			
4	,877	8,768	60,897			
5	,861	8,615	69,512			
6	,696	6,961	76,472			
7	,653	6,535	83,007			
8	,632	6,324	89,331			
9	,556	5,561	94,892			
10	,511	5,108	100,000			

Método de extracción: Análisis de Componentes principales.

Anexo 3

Análisis Factorial de las variables académicas con 9 variables

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,784
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	579,700
	Gl	36
	Sig.	,000

Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,755	30,610	30,610	2,755	30,610	30,610
2	1,209	13,432	44,042	1,209	13,432	44,042
3	,989	10,985	55,027			
4	,866	9,618	64,645			
5	,723	8,038	72,683			
6	,693	7,698	80,381			
7	,638	7,088	87,470			
8	,613	6,807	94,276			
9	,515	5,724	100,000			

Método de extracción: Análisis de Componentes principales.

Anexo 4

Análisis Factorial de las variables académicas con 8 variables

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,774
Chi-cuadrado aproximado		500,548
Prueba de esfericidad de Bartlett	gl	28
	Sig.	,000

Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,606	32,572	32,572	2,606	32,572	32,572
2	1,120	13,999	46,572	1,120	13,999	46,572
3	,969	12,107	58,679			
4	,774	9,673	68,352			
5	,701	8,765	77,117			
6	,690	8,620	85,737			
7	,624	7,806	93,544			
8	,517	6,456	100,000			

Método de extracción: Análisis de Componentes principales.

Anexo 5

Análisis factorial para las variables de entretenimiento con 5 variables

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,650
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	461,607
	gl	10
	Sig.	,000

Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,229	44,578	44,578	2,229	44,578	44,578
2	,970	19,393	63,971			
3	,863	17,268	81,240			
4	,613	12,261	93,501			
5	,325	6,499	100,000			

Método de extracción: Análisis de Componentes principales.

Anexo 6

Análisis factorial para las variables de entretenimiento con 4 variables

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,625
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	386,911
	gl	6
	Sig.	,000

Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,056	51,394	51,394	2,056	51,394	51,394
2	,889	22,218	73,612			
3	,727	18,167	91,779			
4	,329	8,221	100,000			

Método de extracción: Análisis de Componentes principales.

Anexo 7

Análisis factorial para las variables de entretenimiento con 3 variables

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,563
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	319,649
	gl	3
	Sig.	,000

Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	1,839	61,306	61,306	1,839	61,306	61,306
2	,831	27,686	88,992			
3	,330	11,008	100,000			

Método de extracción: Análisis de Componentes principales.

Anexo 8

Análisis clúster (conglomerados) para generar 2 grupos, según el uso de Internet en actividades académicas

Centros de los conglomerados finales

	Conglomerado	
	1	2
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	4,79	8,06
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	5,58	10,61
¿Aproximadamente cuántos videos académicos mira en YouTube cada mes?	4,47	6,04
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	4,05	7,53
¿Aproximadamente cuántas horas busca información académica en Internet cada mes?	7,80	38,19

ANOVA

	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	731,731	1	31,811	468	23,003	,000
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	1729,457	1	41,333	468	41,842	,000
¿Aproximadamente cuántos videos académicos mira en YouTube cada mes?	168,678	1	25,203	468	6,693	,010
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	829,373	1	21,461	468	38,645	,000
¿Aproximadamente cuántas horas busca información académica en Internet cada mes?	63135,861	1	57,101	468	1105,691	,000

Número de casos en cada conglomerado

Conglomerado	1	387,000
	2	83,000
Válidos		470,000
Perdidos		,000

Anexo 9

Análisis clúster (conglomerados) para generar 3 grupos, según el uso de Internet en actividades académicas

Centros de los conglomerados finales

	Conglomerado		
	1	2	3
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	3,65	7,37	8,97
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	4,03	10,39	11,24
¿Aproximadamente cuántos videos académicos mira en YouTube cada mes?	3,89	4,88	6,83
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	3,29	5,83	7,69
¿Aproximadamente cuántas horas busca información académica en Internet cada mes?	5,49	48,44	20,51

ANOVA

	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	1330,044	2	27,749	467	47,931	,000
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	2627,983	2	33,870	467	77,591	,000
¿Aproximadamente cuántos videos académicos mira en YouTube cada mes?	380,524	2	23,989	467	15,863	,000
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	881,226	2	19,509	467	45,170	,000
¿Aproximadamente cuántas horas busca información académica en Internet cada mes?	37846,889	2	30,332	467	1247,734	,000

Número de casos en cada conglomerado

Conglomerado	1	306,000
	2	41,000
	3	123,000
Válidos		470,000
Perdidos		,000

Anexo 10

Análisis clúster (conglomerados) para generar 4 grupos, según el uso de Internet en actividades académicas

Centros de los conglomerados finales

	Conglomerado			
	1	2	3	4
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	3,19	7,37	12,04	6,01
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	3,41	10,39	15,30	7,32
¿Aproximadamente cuántos videos académicos mira en YouTube cada mes?	3,38	4,88	9,21	5,46
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	3,22	5,83	5,85	8,10
¿Aproximadamente cuántas horas busca información académica en Internet cada mes?	5,20	48,44	11,97	24,05

ANOVA

	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	1560,917	3	23,468	466	66,511	,000
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	2941,969	3	26,282	466	111,940	,000
¿Aproximadamente cuántos videos académicos mira en YouTube cada mes?	658,023	3	21,437	466	30,695	,000
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	557,941	3	19,741	466	28,263	,000
¿Aproximadamente cuántas horas busca información académica en Internet cada mes?	26052,621	3	25,110	466	1037,546	,000

Número de casos en cada conglomerado

Conglomerado	1	279,000
	2	41,000
	3	71,000
	4	79,000
Válidos		470,000
Perdidos		,000

Anexo 11

Análisis clúster (conglomerados) para generar 5 grupos, según el uso de Internet en actividades académicas

Centros de los conglomerados finales

	Conglomerado				
	1	2	3	4	5
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	9,62	5,84	3,16	4,62	20,96
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	15,06	9,89	3,30	6,82	12,83
¿Aproximadamente cuántos videos académicos mira en YouTube cada mes?	9,71	5,08	3,25	5,49	5,39
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	6,03	6,13	3,14	8,14	5,22
¿Aproximadamente cuántas horas busca información académica en Internet cada mes?	10,11	49,11	5,08	23,15	26,04

ANOVA

	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	2033,556	4	16,096	465	126,336	,000
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	2226,757	4	26,164	465	85,109	,000
¿Aproximadamente cuántos videos académicos mira en YouTube cada mes?	565,870	4	20,861	465	27,126	,000
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	429,250	4	19,691	465	21,800	,000
¿Aproximadamente cuántas horas busca información académica en Internet cada mes?	19627,890	4	24,403	465	804,316	,000

Número de casos en cada conglomerado

Conglomerado	1	65,000
	2	38,000
	3	271,000
	4	73,000
	5	23,000
Válidos		470,000
Perdidos		,000

Anexo 12

Análisis discriminante 2 grupos, para la determinación del uso de Internet en actividades académicas

Resumen del proceso de clasificación

Procesados		470
Excluidos	Código de grupo perdido o fuera de rango	0
	Perdida al menos una variable discriminante	0
Usados en los resultados		470

Probabilidades previas para los grupos

Número inicial de casos	Previas	Casos utilizados en el análisis	
		No ponderados	Ponderados
1	,500	387	387,000
2	,500	83	83,000
Total	1,000	470	470,000

Resultados de la clasificación^a

Número inicial de casos		Grupo de pertenencia pronosticado		Total	
		1	2		
Original	Recuento	1	384	3	387
		2	3	80	83
	%	1	99,2	,8	100,0
		2	3,6	96,4	100,0

a. Clasificados correctamente el 98,7% de los casos agrupados originales.

Anexo 13

Análisis discriminante 3 grupos, para la determinación del uso de Internet en actividades académicas

Resumen del proceso de clasificación

Procesados		470
Excluidos	Código de grupo perdido o fuera de rango	0
	Perdida al menos una variable discriminante	0
Usados en los resultados		470

Probabilidades previas para los grupos

Número inicial de casos	Previas	Casos utilizados en el análisis	
		No ponderados	Ponderados
1	,333	306	306,000
2	,333	41	41,000
3	,333	123	123,000
Total	1,000	470	470,000

Resultados de la clasificación^a

Número inicial de casos		Grupo de pertenencia pronosticado			Total	
		1	2	3		
Original	Recuento	1	301	0	5	306
		2	0	41	0	41
		3	3	0	120	123
	%	1	98,4	,0	1,6	100,0
		2	,0	100,0	,0	100,0
		3	2,4	,0	97,6	100,0

a. Clasificados correctamente el 98,3% de los casos agrupados originales.

Anexo 14

Análisis discriminante 4 grupos, para la determinación del uso de Internet en actividades académicas

Resumen del proceso de clasificación

Procesados		470
Excluidos	Código de grupo perdido o fuera de rango	0
	Perdida al menos una variable discriminante	0
Usados en los resultados		470

Probabilidades previas para los grupos

Número inicial de casos	Previas	Casos utilizados en el análisis	
		No ponderados	Ponderados
1	,250	279	279,000
2	,250	41	41,000
3	,250	71	71,000
4	,250	79	79,000
Total	1,000	470	470,000

Resultados de la clasificación^a

Número inicial de casos		Grupo de pertenencia pronosticado				Total	
		1	2	3	4		
Original	Recuento	1	275	0	2	2	279
		2	0	40	0	1	41
		3	2	0	67	2	71
		4	0	0	1	78	79
	%	1	98,6	,0	,7	,7	100,0
		2	,0	97,6	,0	2,4	100,0
		3	2,8	,0	94,4	2,8	100,0
		4	,0	,0	1,3	98,7	100,0

a. Clasificados correctamente el 97,9% de los casos agrupados originales.

Anexo 15

Análisis discriminante 5 grupos, para la determinación del uso de Internet en actividades académicas

Resumen del proceso de clasificación

Procesados		470
Excluidos	Código de grupo perdido o fuera de rango	0
	Perdida al menos una variable discriminante	0
Usados en los resultados		470

Probabilidades previas para los grupos

Número inicial de casos	Previas	Casos utilizados en el análisis	
		No ponderados	Ponderados
1	,200	65	65,000
2	,200	38	38,000
3	,200	271	271,000
4	,200	73	73,000
5	,200	23	23,000
Total	1,000	470	470,000

Resultados de la clasificación^a

Número inicial de casos		Grupo de pertenencia pronosticado					Total	
		1	2	3	4	5		
Original	Recuento	1	62	0	1	1	1	65
		2	0	38	0	0	0	38
		3	9	0	259	3	0	271
		4	0	0	0	73	0	73
		5	0	0	0	2	21	23
	%	1	95,4	,0	1,5	1,5	1,5	100,0
		2	,0	100,0	,0	,0	,0	100,0
		3	3,3	,0	95,6	1,1	,0	100,0
		4	,0	,0	,0	100,0	,0	100,0
		5	,0	,0	,0	8,7	91,3	100,0

a. Clasificados correctamente el 96,4% de los casos agrupados originales.

Anexo 16

Clasificación de los grupos del Uso de Internet en actividades académicas por género
(tablas de contingencia)

Resumen del procesamiento de los casos

	Casos					
	Válidos		Perdidos		Total	
	N	Porcentaje	N	Porcentaje	N	Porcentaje
¿Cuál es su género? * Uso de Internet en actividades académicas	470	100,0%	0	,0%	470	100,0%

Tabla de contingencia ¿Cuál es su género? * Uso de Internet en actividades académicas

			Uso de Internet en actividades académicas		Total
			1	2	
¿Cuál es su género?	Hombre	Recuento	203	46	249
		% dentro de ¿Cuál es su género?	81,5%	18,5%	100,0%
		% dentro de Uso de Internet en actividades académicas	52,5%	55,4%	53,0%
		% del total	43,2%	9,8%	53,0%
	Mujer	Recuento	184	37	221
		% dentro de ¿Cuál es su género?	83,3%	16,7%	100,0%
		% dentro de Uso de Internet en actividades académicas	47,5%	44,6%	47,0%
		% del total	39,1%	7,9%	47,0%
Total	Recuento	387	83	470	
	% dentro de ¿Cuál es su género?	82,3%	17,7%	100,0%	
	% dentro de Uso de Internet en actividades académicas	100,0%	100,0%	100,0%	
	% del total	82,3%	17,7%	100,0%	

Medidas simétricas

		Valor	Error típ. asint. ^a	T aproximada ^b	Sig. aproximada
Intervalo por intervalo	R de Pearson	-,023	,046	-,490	,624 ^c
Ordinal por ordinal	Correlación de Spearman	-,023	,046	-,490	,624 ^c
N de casos válidos		470			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.

c. Basada en la aproximación normal.

Anexo 17

Determinación de la varianza de edad para los dos grupos de Uso de Internet en actividades académicas

Estadísticos de grupo

Uso de Internet en actividades académicas		N	Media	Desviación típ.	Error típ. de la media
¿Cuál es su edad?	1	387	23,33	4,043	,205
	2	83	23,45	3,855	,423

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
									Inferior	Superior
¿Cuál es su edad?	Se han asumido varianzas iguales	1,662	,198	-,232	468	,817	-,112	,485	-1,066	,841
	No se han asumido varianzas iguales			-,239	123,780	,811	-,112	,470	-1,043	,819

Anexo 18

Análisis clúster (conglomerados) para generar 2 grupos, según el uso de Internet para entretenimiento

Centros de los conglomerados finales

	Conglomerado	
	1	2
¿Aproximadamente cuántas horas a la semana chatea por diversión?	6,40	17,51
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	6,84	23,76
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	1,84	6,67
¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	3,83	6,21
¿Aproximadamente cuántos videos para entretenimiento mira en YouTube cada semana?	6,67	10,89

ANOVA

	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas horas a la semana chatea por diversión?	11212,214	1	25,449	468	440,579	,000
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	25997,381	1	29,918	468	868,946	,000
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	2111,751	1	31,378	468	67,300	,000
¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	514,998	1	12,858	468	40,053	,000
¿Aproximadamente cuántos videos para entretenimiento mira en YouTube cada semana?	1610,934	1	47,510	468	33,907	,000

Las pruebas F sólo se deben utilizar con una finalidad descriptiva puesto que los conglomerados han sido elegidos para maximizar las diferencias entre los casos en diferentes conglomerados. Los niveles críticos no son corregidos, por lo que no pueden interpretarse como pruebas de la hipótesis de que los centros de los conglomerados son iguales.

Número de casos en cada conglomerado

Conglomerado	1	347,000
	2	123,000
Válidos		470,000
Perdidos		,000

Anexo 19

Análisis clúster (conglomerados) para generar 3 grupos, según el uso de Internet para entretenimiento

Centros de los conglomerados finales

	Conglomerado		
	1	2	3
¿Aproximadamente cuántas horas a la semana chatea por diversión?	6,35	14,00	17,96
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	6,77	21,58	23,70
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	1,72	24,38	2,76
¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	3,86	7,21	5,77
¿Aproximadamente cuántos videos para entretenimiento mira en YouTube cada semana?	6,69	12,88	10,17

ANOVA

	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas horas a la semana chatea por diversión?	5656,413	2	25,288	467	223,681	,000
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	12777,398	2	30,930	467	413,106	,000
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	5763,506	2	11,284	467	510,764	,000
¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	241,394	2	12,954	467	18,634	,000
¿Aproximadamente cuántos videos para entretenimiento mira en YouTube cada semana?	812,443	2	47,582	467	17,075	,000

Las pruebas F sólo se deben utilizar con una finalidad descriptiva puesto que los conglomerados han sido elegidos para maximizar las diferencias entre los casos en diferentes conglomerados. Los niveles críticos no son corregidos, por lo que no pueden interpretarse como pruebas de la hipótesis de que los centros de los conglomerados son iguales.

Número de casos en cada conglomerado

Conglomerado	1	342,000
	2	24,000
	3	104,000
Válidos		470,000
Perdidos		,000

Anexo 20

Análisis clúster (conglomerados) para generar 4 grupos, según el uso de Internet para entretenimiento

Centros de los conglomerados finales

	Conglomerado			
	1	2	3	4
¿Aproximadamente cuántas horas a la semana chatea por diversión?	5,83	9,49	16,08	17,96
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	6,25	9,68	25,04	24,38
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	1,67	2,86	23,54	2,28
¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	3,10	7,30	6,71	5,43
¿Aproximadamente cuántos videos para entretenimiento mira en YouTube cada semana?	4,12	18,23	11,33	8,58

ANOVA

	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas horas a la semana chatea por diversión?	3718,686	3	25,679	466	144,817	,000
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	9049,228	3	27,578	466	328,129	,000
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	3553,313	3	13,169	466	269,825	,000
¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	456,058	3	11,082	466	41,152	,000
¿Aproximadamente cuántos videos para entretenimiento mira en YouTube cada semana?	4296,638	3	23,510	466	182,758	,000

Las pruebas F sólo se deben utilizar con una finalidad descriptiva puesto que los conglomerados han sido elegidos para maximizar las diferencias entre los casos en diferentes conglomerados. Los niveles críticos no son corregidos, por lo que no pueden interpretarse como pruebas de la hipótesis de que los centros de los conglomerados son iguales.

Número de casos en cada conglomerado

Conglomerado	1	275,000
	2	81,000
	3	24,000
	4	90,000
Válidos		470,000
Perdidos		,000

Anexo 21

Análisis clúster (conglomerados) para generar 5 grupos, según el uso de Internet para entretenimiento

Centros de los conglomerados finales

	Conglomerado				
	1	2	3	4	5
¿Aproximadamente cuántas horas a la semana chatea por diversión?	9,17	10,04	14,81	19,58	3,28
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	9,31	10,70	23,86	26,92	3,86
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	2,50	2,35	24,86	3,06	1,25
¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	6,62	4,96	7,14	5,43	2,38
¿Aproximadamente cuántos videos para entretenimiento mira en YouTube cada semana?	21,28	6,18	12,05	9,11	3,31

ANOVA

	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas horas a la semana chatea por diversión?	3534,054	4	19,325	465	182,876	,000
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	7505,464	4	21,457	465	349,799	,000
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	2649,964	4	13,327	465	198,849	,000
¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	306,100	4	11,415	465	26,815	,000
¿Aproximadamente cuántos videos para entretenimiento mira en YouTube cada semana?	3670,547	4	19,706	465	186,264	,000

Las pruebas F sólo se deben utilizar con una finalidad descriptiva puesto que los conglomerados han sido elegidos para maximizar las diferencias entre los casos en diferentes conglomerados. Los niveles críticos no son corregidos, por lo que no pueden interpretarse como pruebas de la hipótesis de que los centros de los conglomerados son iguales.

Número de casos en cada conglomerado

Conglomerado	1	58,000
	2	159,000
	3	21,000
	4	72,000
	5	160,000
Válidos		470,000
Perdidos		,000

Anexo 22

Análisis discriminante para determinar el nivel de exactitud en una clasificación de 2 grupos, para la determinación del uso de Internet para el entretenimiento

Resumen del proceso de clasificación

Procesados		470
Excluidos	Código de grupo perdido o fuera de rango	0
	Perdida al menos una variable discriminante	0
Usados en los resultados		470

Probabilidades previas para los grupos

Número inicial de casos	Previas	Casos utilizados en el análisis	
		No ponderados	Ponderados
1	,500	347	347,000
2	,500	123	123,000
Total	1,000	470	470,000

Resultados de la clasificación^a

Número inicial de casos			Grupo de pertenencia pronosticado		Total
			1	2	
Original	Recuento	1	343	4	347
		2	1	122	123
	%	1	98,8	1,2	100,0
		2	,8	99,2	100,0

a. Clasificados correctamente el 98,9% de los casos agrupados originales.

Anexo 23

Análisis discriminante para determinar el nivel de exactitud en una clasificación de 3 grupos, para la determinación del uso de Internet para el entretenimiento

Resumen del proceso de clasificación

Procesados		470
Excluidos	Código de grupo perdido o fuera de rango	0
	Perdida al menos una variable discriminante	0
Usados en los resultados		470

Probabilidades previas para los grupos

Número inicial de casos	Previas	Casos utilizados en el análisis	
		No ponderados	Ponderados
1	,333	342	342,000
2	,333	24	24,000
3	,333	104	104,000
Total	1,000	470	470,000

Resultados de la clasificación^a

Número inicial de casos			Grupo de pertenencia pronosticado			Total
			1	2	3	
Original	Recuento	1	337	2	3	342
		2	0	24	0	24
		3	2	0	102	104
	%	1	98,5	,6	,9	100,0
		2	,0	100,0	,0	100,0
		3	1,9	,0	98,1	100,0

a. Clasificados correctamente el 98,5% de los casos agrupados originales.

Anexo 24

Análisis discriminante para determinar el nivel de exactitud en una clasificación de 4 grupos, para la determinación del uso de Internet para el entretenimiento

Resumen del proceso de clasificación

Procesados		470
Excluidos	Código de grupo perdido o fuera de rango	0
	Perdida al menos una variable discriminante	0
Usados en los resultados		470

Probabilidades previas para los grupos

Número inicial de casos	Previas	Casos utilizados en el análisis	
		No ponderados	Ponderados
1	,250	275	275,000
2	,250	81	81,000
3	,250	24	24,000
4	,250	90	90,000
Total	1,000	470	470,000

Resultados de la clasificación^a

Número inicial de casos			Grupo de pertenencia pronosticado				Total
			1	2	3	4	
Original	Recuento	1	269	3	2	1	275
		2	2	77	1	1	81
		3	0	0	24	0	24
		4	1	5	0	84	90
	%	1	97,8	1,1	,7	,4	100,0
		2	2,5	95,1	1,2	1,2	100,0
		3	,0	,0	100,0	,0	100,0
		4	1,1	5,6	,0	93,3	100,0

a. Clasificados correctamente el 96,6% de los casos agrupados originales.

Anexo 25

Análisis discriminante para determinar el nivel de exactitud en una clasificación de 5 grupos, para la determinación del uso de Internet para el entretenimiento

Resumen del proceso de clasificación

Procesados		470
Excluidos	Código de grupo perdido o fuera de rango	0
	Perdida al menos una variable discriminante	0
Usados en los resultados		470

Probabilidades previas para los grupos

Número inicial de casos	Previas	Casos utilizados en el análisis	
		No ponderados	Ponderados
1	,200	58	58,000
2	,200	159	159,000
3	,200	21	21,000
4	,200	72	72,000
5	,200	160	160,000
Total	1,000	470	470,000

Resultados de la clasificación^a

Número inicial de casos		Grupo de pertenencia pronosticado					Total	
		1	2	3	4	5		
Original	Recuento	1	57	0	1	0	0	58
		2	1	157	0	0	1	159
		3	0	0	21	0	0	21
		4	0	0	0	72	0	72
		5	0	0	1	0	159	160
	%	1	98,3	,0	1,7	,0	,0	100,0
		2	,6	98,7	,0	,0	,6	100,0
		3	,0	,0	100,0	,0	,0	100,0
		4	,0	,0	,0	100,0	,0	100,0
		5	,0	,0	,6	,0	99,4	100,0

a. Clasificados correctamente el 99,1% de los casos agrupados originales.

Anexo 26

Clasificación del género según el uso de Internet en actividades de entretenimiento

Resumen del procesamiento de los casos

	Casos					
	Válidos		Perdidos		Total	
	N	Porcentaje	N	Porcentaje	N	Porcentaje
¿Cuál es su género? * Usos de Internet en entretenimiento	470	100,0%	0	0,0%	470	100,0%

Tabla de contingencia ¿Cuál es su género? * Usos de Internet en entretenimiento

		Usos de Internet en entretenimiento		Total	
		1	2		
¿Cuál es su género?	Hombre	Recuento	184	65	249
		% dentro de ¿Cuál es su género?	73,9%	26,1%	100,0%
		% dentro de Usos de Internet en entretenimiento	53,0%	52,8%	53,0%
	Mujer	Recuento	163	58	221
		% dentro de ¿Cuál es su género?	73,8%	26,2%	100,0%
		% dentro de Usos de Internet en entretenimiento	47,0%	47,2%	47,0%
Total	Recuento	347	123	470	
	% dentro de ¿Cuál es su género?	73,8%	26,2%	100,0%	
	% dentro de Usos de Internet en entretenimiento	100,0%	100,0%	100,0%	

Medidas simétricas

		Valor	Error típ. asint. ^a	T aproximada ^b	Sig. aproximada
Nominal por nominal	Coefficiente de contingencia	,002			,973
Intervalo por intervalo	R de Pearson	,002	,046	,034	,973 ^c
Ordinal por ordinal	Correlación de Spearman	,002	,046	,034	,973 ^c
N de casos válidos		470			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.

c. Basada en la aproximación normal.

Anexo 27

Relación entre la edad y el uso de Internet en entretenimiento

Estadísticos de grupo

	Usos de Internet en entretenimiento	N	Media	Desviación típ.	Error típ. de la media
¿Cuál es su edad?	1	347	23,54	4,089	,219
	2	123	22,81	3,727	,336

Prueba de muestras independientes

	Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias							
	F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia		
								Inferior	Superior	
¿Cuál es su edad?	1,833	,176	1,744	468	,082	,732	,419	-,093	1,556	
Se han asumido varianzas iguales			1,823	233,315	,070	,732	,401	-,059	1,522	
No se han asumido varianzas iguales										

Anexo 28
Regresión logística binomial hipótesis 1

Codificación de la variable dependiente

Valor original	Valor interno
Preocupado	0
Despreocupado	1

Codificaciones de variables categóricas

	Frecuencia	Codificación de parámetros				
		(1)	(2)	(3)	(4)	
Los ingresos mensuales de su familia son de:	Hasta 350 dólares	60	1,000	,000	,000	,000
	Hasta 600 dólares	97	,000	1,000	,000	,000
	Hasta 1000 dólares	112	,000	,000	1,000	,000
	Hasta 1.500 dólares	90	,000	,000	,000	1,000
	Más de 1.500 dólares	111	,000	,000	,000	,000

Pruebas omnibus sobre los coeficientes del modelo

	Chi cuadrado	Gl	Sig.
Paso	1,769	4	,778
Paso 1 Bloque	1,769	4	,778
Modelo	1,769	4	,778

Resumen del modelo

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	436,452 ^a	,004	,006

a. La estimación ha finalizado en el número de iteración 4 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,000	3	1,000

Tabla de contingencias para la prueba de Hosmer y Lemeshow

	Usos de Internet en actividades académicas = Preocupado		Usos de Internet en actividades académicas = Despreocupado		Total
	Observado	Esperado	Observado	Esperado	
	1	76	76,000	14	
2	93	93,000	18	18,000	111
Paso 1	3	93,000	19	19,000	112
4	79	79,000	18	18,000	97
5	46	46,000	14	14,000	60

Anexo 29

Regresión logística binomial hipótesis 2

Codificación de la variable dependiente

Valor original	Valor interno
Activo	0
Pasivo	1

Codificaciones de variables categóricas

	Frecuencia	Codificación de parámetros				
		(1)	(2)	(3)	(4)	
Los ingresos mensuales de su familia son de:	Hasta 350 dólares	60	1,000	,000	,000	,000
	Hasta 600 dólares	97	,000	1,000	,000	,000
	Hasta 1000 dólares	112	,000	,000	1,000	,000
	Hasta 1.500 dólares	90	,000	,000	,000	1,000
	Más de 1.500 dólares	111	,000	,000	,000	,000

Pruebas omnibus sobre los coeficientes del modelo

	Chi cuadrado	gl	Sig.
Paso	11,324	4	,023
Paso 1 Bloque	11,324	4	,023
Modelo	11,324	4	,023

Resumen del modelo

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	529,016 ^a	,024	,035

a. La estimación ha finalizado en el número de iteración 4 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,000	3	1,000

Tabla de contingencias para la prueba de Hosmer y Lemeshow

	Usos de Internet en entretenimiento = Activo		Usos de Internet en entretenimiento = Pasivo		Total	
	Observado	Esperado	Observado	Esperado		
	Paso 1	1	80	80,000		17
	2	46	46,000	14	14,000	60
	3	83	83,000	28	28,000	111
	4	83	83,000	29	29,000	112
	5	55	55,000	35	35,000	90

Anexo 30
Regresión logística binomial hipótesis 3

Codificación de la variable dependiente

Valor original	Valor interno
Destacado	0
Común	1

Codificaciones de variables categóricas

		Frecuencia	Codificación de parámetros
			(1)
Usos de Internet en actividades académicas	Preocupado	387	1,000
	Despreocupado	83	,000

Pruebas omnibus sobre los coeficientes del modelo

		Chi cuadrado	gl	Sig.
Paso		4,552	1	,033
Paso 1	Bloque	4,552	1	,033
	Modelo	4,552	1	,033

Resumen del modelo

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	579,466 ^a	,010	,014

a. La estimación ha finalizado en el número de iteración 4 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,000	0	1,000

Tabla de contingencias para la prueba de Hosmer y Lemeshow

	rend_acad = Destacado		rend_acad = Común		Total	
	Observado	Esperado	Observado	Esperado		
Paso 1	1	65	65,000	18	18,000	83
	2	258	258,000	129	129,000	387

Anexo 31

Regresión logística para comprobación de hipótesis 4

Codificación de la variable dependiente

Valor original	Valor interno
Destacado	0
Común	1

Codificaciones de variables categóricas

		Frecuencia	Codificación de parámetros
			(1)
Usos de Internet en entretenimiento	Activo	347	1,000
	Pasivo	123	,000

Pruebas omnibus sobre los coeficientes del modelo

		Chi cuadrado	gl	Sig.
Paso		6,623	1	,010
Paso 1	Bloque	6,623	1	,010
	Modelo	6,623	1	,010

Resumen del modelo

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	577,395 ^a	,014	,020

a. La estimación ha finalizado en el número de iteración 3 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,000	0	1.000

Tabla de contingencias para la prueba de Hosmer y Lemeshow

		rend_acad = Destacado		rend_acad = Común		Total
		Observado	Esperado	Observado	Esperado	
Paso 1	1	250	250,000	97	97,000	347
	2	73	73,000	50	50,000	123

Anexo 32:

Descripción del Proyecto de titulación “Educación Virtual en el Ecuador”

Universidad Técnica Particular de Loja

TITULACIÓN DE INGENIERÍA EN SISTEMAS INFORMÁTICOS Y COMPUTACIÓN

Guía para elaboración de Propuestas de Proyectos de Fin de Titulación

Información General del Proyecto

Título del proyecto:	Educación virtual en Ecuador
Duración:	9 meses
Propuesto por:	Equipo:

Juan Carlos Torres Díaz
Carlos Correa Granda
Héctor Gómez
Inés Jara
Diego Alvarado
Daysi Karina García
Lorena Condolo
Martha Agila
Pablo Torres
Rosario Requena
Paola Andrade

Docente Investigador: Juan Carlos Torres Díaz (jctorres@utpl.edu.ec)

Línea de Investigación: Tecnología aplicada a la educación

Perfil Requerido del estudiante	35 Egresados de Ingeniería en Informática
---------------------------------	---

Propósito / Descripción

El proyecto busca levantar información de las universidades ecuatorianas de las categorías A, B y C, esta información permitirá contar con una línea de base respecto al uso de la tecnología en el sistema educativo del país, esta información es fundamental para la toma de decisiones que permitan optimizar el uso de la tecnología y maximizar sus efectos en el rendimiento académico de los estudiantes.

Se levantará información referente a como los estudiantes utilizan la tecnología en sus actividades académicas y de entretenimiento. Se buscará patrones de uso de la tecnología y se categorizará a los estudiantes en grupos homogéneos; así también se buscará la existencia de relaciones entre los usos de la tecnología y el éxito académico que alcanzan los estudiantes lo que permitirá determinar el real efecto de las herramientas informáticas dentro del sistema educativo.

Un tema sensible que requiere de un análisis sustentado es el que tiene que ver con el uso ético de la tecnología, esto incluye el plagio de información en la red y las implicaciones que tendría sobre los niveles de aprendizaje y el éxito académico.

Competencias a desarrollar

Levantar, tabular y organizar datos para la toma de decisiones

Inferir y generar información a partir de datos organizados

Componentes:

Componente	Descripción
Diseño de la investigación	En este componente se van a definir todos los aspectos metodológicos que se requieren previo inicio del trabajo, se definen aquí instituciones, tamaños de muestras, cuestionarios, procedimientos estadísticos de análisis. Este trabajo está a cargo del proponente del proyecto y será dado a conocer a los participantes.
Inducción a tesis	Es una etapa de formación en la que se da a conocer de forma detallada la metodología a emplear, el procedimiento, aspectos legales, plazos y entregables.
Levantamiento y organización de datos	En este componente, los participantes encuestan a los estudiantes y profesores de las universidades, digitan los datos y realizan un análisis preliminar de los mismos.
Tratamiento y análisis de datos	En esta etapa se aplican procedimientos estadísticos a fin de obtener, entre otros resultados, categorizaciones, correlaciones, modelos de regresión. Con esos insumos se procede a documentar las relaciones y hallazgos.
Generación de informe	Es la etapa final, consiste en documentar el trabajo y consta de los siguientes apartados: Introducción Metodología Resultados Discusión de resultados Conclusiones Referencias

Estrategia o Metodología de desarrollo (Opcional)

Como parte inicial, la UTPL a través del director del proyecto diseñará el proyecto en sus aspectos metodológicos y organizativos. Posteriormente se procederá a instruir a los estudiantes y a poner en práctica los siguientes aspectos:

Aspecto 1: Autorización y levantamiento de datos para una muestra representativa en cada institución seleccionada

Aspecto 2: Utilización del cuestionario “usos de tecnología en las universidades” basado en los utilizados en los proyectos PIC, DLINHE, ECUADOR

Aspecto 3: Aplicación de análisis multivariante para obtener categorizaciones respecto a los usos de Internet para la educación y para el entretenimiento.

Aspecto 4: Aplicación de regresiones (lineales, logísticas) para obtener modelos explicativos y/o predictivos que expliquen las distintas realidades.

Aspecto 5: Levantamiento de información cualitativa utilizando las técnicas: entrevista semi-estructurada, cuestionarios y búsqueda de información institucional en la red.

Resultados esperados

Informe de investigación con los apartados siguientes:

Introducción: que se trata el trabajo / objetivos / preguntas de investigación /importancia del trabajo / componentes del trabajo

Marco teórico: revisión teórica de por lo menos 25 autores que traten los temas de brecha digital/usos de Internet/rendimiento académico/

Metodología: provista por la UTPL

Resultados: Descripción de los hallazgos

Discusión de los resultados:

Conclusiones.

Cronograma

En función de cada componente, indicar la duración estimada.

Componente	Tiempo
Diseño de la investigación	1 mes
Inducción y formación de los investigadores	1 mes
Levantamiento y organización de datos	2 meses
Tratamiento y análisis de datos	1 meses
Generación de informe	2 mes
Presentación y trámites legales	1 mes

Bibliografía

Torres-Díaz, J. C., Morocho, M., & Guamán, J. (2010). La educación virtual en Ecuador. In C. Rama & J. L. Pardo (Eds.), *La educación superior a distancia: Miradas diversas desde Iberoamérica* (pp. 76–91). Madrid: INTEVED.

Torres-Díaz, J. C., & Infante-Moro, A. (2011). Desigualdad digital en la universidad: Incidencia de los ingresos en los usos de Internet en Ecuador. *Comunicar*, 37(2). doi:10.3916/C37-2011-02-08

Anexo 1: Listado de universidades de categorías A, B y C

Universidades Categoría A

1. Escuela Politécnica Nacional
2. Escuela Superior Politécnica del Litoral
3. Universidad San Francisco de Quito

Universidades Categoría B

4. Escuela Superior Politécnica de Chimborazo
5. Pontificia Universidad Católica del Ecuador
6. Universidad Casa Grande Universidad Católica
7. Santiago de Guayaquil
8. Universidad Central del Ecuador
9. Universidad de Cuenca
10. Universidad del Azuay
11. Universidad Estatal de Milagro
12. Universidad Nacional de Loja
13. Universidad Particular Internacional Sek
14. Universidad Politécnica Salesiana
15. Universidad Técnica de Ambato
16. Universidad Técnica del Norte
17. Universidad Técnica Estatal de Quevedo
18. Universidad Técnica Particular de Loja
19. Universidad Tecnológica empresarial de Guayaquil
20. Universidad Tecnológica Equinoccial
21. Universidad Tecnológica Indoamérica

Universidades categoría C

22. Escuela Superior Politécnica agropecuaria de Manabí
23. Universidad de Especialidades Turísticas
24. Universidad de las Américas
25. Universidad del Pacífico – Escuela de Negocios
26. Universidad Estatal del Bolívar
27. Universidad Internacional del Ecuador
28. Universidad Laica Vicente Rocafuerte del Ecuador
29. Universidad Metropolitana
30. Universidad nacional de Chimborazo
31. Universidades de Especialidades Espíritu Santo
32. Universidad Regional Autónoma de los Andes
33. Universidad Técnica de Babahoyo
34. Universidad Técnica de Cotopaxi
35. Universidad Técnica de Israel

Anexo 33

Autorización para levantamiento de información UMET



Quito, 20 de Octubre del 2014.

Ph. D
JUAN PABLO SUAREZ CHACÓN
Director General de Investigación
Universidad Técnica Particular de Loja
Loja.-

En relación a su **oficio Nro. 390-DGITT-UTPL** con fecha 08 de septiembre de 2014, tengo a bien comunicar que se autoriza la aplicación de quinientas encuestas a estudiantes de la Universidad Metropolitana y una entrevista a un Directivo de la misma, con el objeto de apoyar al Proyecto de Investigación que desarrolla la UTPL, las mismas que serán realizadas por el alumno **Juan Paúl Jiménez Gaona** con cédula de identidad **Nº 1104695489**.

Reitero mis sentidos de consideración.

Cordialmente,

A handwritten signature in blue ink, appearing to read 'JLB', is written over a light blue circular stamp.

ING. JORGE LUIS BASANTES
COORDINADOR ACADÉMICO UMET



DIRECCION
ACADEMICA

www.umet.edu.ec
Dirección: Av La Coruna N26-95
Telf.: (02) 222-9544
Quito - Ecuador