

**ANÁLISIS DE LOS FACTORES ASOCIADOS AL RIESGO DE DESERCIÓN EN  
LOS ESTUDIANTES DEL PROGRAMA LICENCIATURA EN INFORMÁTICA DE LA  
UNIVERSIDAD DE NARIÑO, PARA LA CONSTRUCCIÓN Y SISTEMATIZACIÓN DE  
UN MODELO ESTADÍSTICO PREDICTIVO**

**MICHELL DANIELA BRAVO MORENO**

**HAILY TATIANA MAYA PIANDA**

UNIVERSIDAD DE NARIÑO  
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA  
FACULTAD DE CIENCIAS EXACTAS Y NATURALES  
PROGRAMA DE LICENCIATURA EN INFORMÁTICA  
SAN JUAN DE PASTO

2018

**ANÁLISIS DE LOS FACTORES ASOCIADOS AL RIESGO DE DESERCIÓN EN  
LOS ESTUDIANTES DEL PROGRAMA LICENCIATURA EN INFORMÁTICA DE LA  
UNIVERSIDAD DE NARIÑO, PARA LA CONSTRUCCIÓN Y SISTEMATIZACIÓN DE  
UN MODELO ESTADÍSTICO PREDICTIVO.**

**MICHEL DANIELA BRAVO MORENO**

**HAILY TATIANA MAYA PIANDA**

Trabajo de grado presentado como requisito parcial para optar por el título de

Licenciadas en Informática

Asesor:

**ARSENIO HIDALGO TROYA**

Magister en Estadística

UNIVERSIDAD DE NARIÑO  
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA  
FACULTAD DE CIENCIAS EXACTAS Y NATURALES  
PROGRAMA DE LICENCIATURA EN INFORMÁTICA  
SAN JUAN DE PASTO

2018

**NOTA DE RESPONSABILIDAD**

Las ideas y conclusiones aportadas en este Trabajo de Grado son Responsabilidad de los autores.

Artículo 1 del Acuerdo No. 324 de octubre 11 de 1966, emanado por el Honorable Consejo Directivo de la Universidad de Nariño.

**Nota de Aceptación**

---

---

---

---

---

---

**Arsenio Hidalgo Troya**

Asesor

---

**Edwin Insuasti**

Jurado

---

**Hernán Abdón García**

Jurado

San Juan de Pasto, 3 de septiembre de 2018

## **AGRADECIMIENTO**

A Dios por la vida y la sabiduría,

A la Universidad de Nariño por abrirnos las puertas al conocimiento,

A los docentes del programa de Licenciatura en Informática por formarnos como profesionales, por su compromiso y dedicación, gracias por incentivarnos a seguir adelante,

Al nuestro Asesor Mg. Arsenio Hidalgo por la colaboración y direccionamiento de este trabajo de grado, por confiar en nuestros saberes y capacidades, por su amabilidad, paciencia y disponibilidad.

A los jurados de esta investigación: Hernán García y Edwin Insuasti por los aportes a este proyecto para su mejoramiento.

## DEDICATORIA

A Dios por permitirme cumplir esta meta, por darme sabiduría y fuerzas en los momentos más difíciles, por ser luz e iluminar mi camino, por impedir que dejara de creer en mí.

A mi más grande inspiración, Valeria y Alejandro quienes dan sentido a mi vida, por ustedes conocí una forma incondicional de amar, gracias mis amores porque he aprendido la profesión más linda del mundo; ser madre.

A mis padres por darme la vida, por ser el mejor ejemplo a seguir y alentarme en todo momento, gracias por todo su afecto.

A Rolando por su apoyo y comprensión, gracias por tus consejos y por acompañarme durante este proceso de realización profesional.

A mis hermanos por estar siempre ahí, por ser compañeros de vida.

*Tatiana Maya*

**DEDICATORIA**

A Dios, por darme fortaleza y sabiduría al emprender el camino del conocimiento,

A mis Padres Humberto y Nelly, por su esfuerzo constante, amor y ejemplo, por enseñarme que la vida es un reto y mostrarme el camino hacia la superación, son mi más grande motivación, este logro se los debo a ustedes.

A Carolina, Dylan y Samir por su compañía en los momentos más turbulentos, por su afecto y cariño, porque son el motivo y las ganas de buscar lo mejor para ustedes.

*Michell Bravo*

## Resumen

La deserción universitaria es un problema constante y profundo que afecta fuertemente la educación superior, produciendo un impacto negativo, tanto en el estudiante, como en su familia y la Institución. De acuerdo con esto, la presente investigación se desarrolla con la finalidad de analizar la asociación que existe entre los diferentes factores relacionados a un estudiante y la decisión de abandonar la carrera.

En primer lugar se hace una caracterización de la población, conformada por los estudiantes del programa de Licenciatura en Informática de la Universidad de Nariño, con el fin de identificar el comportamiento general de la población según los factores personales, académicos, socioeconómicos e institucionales.

Luego, se realiza un análisis bivariado en el cual se incluyen los factores que presuntamente explican la variable deserción, de esta manera se clasifican las variables de acuerdo con el nivel de significancia y se construye el perfil del estudiante desertor.

Posteriormente se hace un análisis multivariado con los resultados obtenidos, se aplica el modelo de regresión logística binaria, para identificar las variables que guardan un alto grado de asociación. Con esta información se obtiene la ecuación que permite calcular la probabilidad de deserción de un estudiante. Finalmente, se sistematiza el modelo por medio de una aplicación para darlo a conocer de una forma más práctica y concreta, de manera que cualquier estudiante o persona interesada pueda realizar el cálculo. De acuerdo con los resultados obtenidos se proponen unas estrategias encaminadas a disminuir los índices de deserción.



**ABSTRACT**

College desertion is a constant and profound problem that affects higher education, it has a negative impact on the student, his family and the institution. According to this, the present investigation is developed with the purpose of analyzing the association that exists between the different factors related to a student and the decision to abandon the career.

In the first place, a characterization of the population is made, formed by the students of the Computer Science degree program of the University of Nariño, in order to identify the general behavior of the population according to personal, academic, socioeconomic and institutional factors.

Then, a bivariate analysis will be carried out, which includes the factors that explain the desertion variable in advance, thus classifying the variables according to the level of significance and constructing the profile of the desert student.

Afterwards, a multivariate analysis is made with the obtained results, the binary logistic regression model is applied, to identify the variables that have a high degree of association. With this information we obtain the equation that allows us to calculate the probability of a student's desertion. Finally, the model is systematized by means of an application to make it known in a more practical and concrete way, in such a way that the interested person can perform the calculation. The agreement with the results was ahead of the strategies aimed at decreasing the dropout rates.

## Tabla de Contenido

	Pág.
<b>1. Aspectos Generales</b> .....	19
1.1 Problema de Investigación.....	19
1.1.1 Planteamiento del Problema.....	19
1.1.2 Formulación del Problema. ....	20
1.2 Objetivos.....	20
1.2.1 Objetivo General. ....	20
1.2.2 Objetivos Específicos.....	20
1.3 Justificación.....	21
<b>2. Marco Referencial</b> .....	23
2.1 Marco Teórico.....	23
2.1.1 Deserción Estudiantil Universitaria. ....	23
2.1.2 Tipos de Deserción Estudiantil. ....	24
2.1.2.1 Deserción precoz. ....	24
2.1.2.2 Deserción temprana. ....	25
2.1.2.3 Deserción tardía.....	25
2.1.2.4 Deserción institucional. ....	25
2.1.2.5 Deserción interna o del programa académico. ....	25
2.1.3 Factores de Deserción. ....	26
2.1.4 Análisis Estadístico. ....	27
2.1.4.1 Análisis Univariado.....	27
2.1.4.2 Análisis Bivariado. ....	27
2.1.4.3 Análisis Multivariado. ....	27
2.1.5 Chi- cuadrado de Pearson.....	28
2.1.6 Odds Ratio.....	30
2.1.7 Regresión Logística.....	32
2.1.8 Regresión logística binaria.....	32
2.1.9 Curva ROC.....	33
2.1.10 Estrategias. ....	34

2.2 Marco Conceptual.....	35
2.3 Marco Legal.....	39
2.4 Antecedentes.....	41
2.5 Marco de Contextual.....	47
<b>3. Procedimiento Metodológico .....</b>	<b>48</b>
3.1 Aspectos investigativos.....	48
3.1.1 Enfoque. ....	48
3.1.2 Alcance.....	49
3.1.3 Diseño. ....	49
3.1.3.1 Operacionalización de Variables.....	49
3.1.4 Población y Muestra.....	52
3.1.5 Métodos de recolección de la información. ....	52
3.2 Diseño Procedimental.....	53
3.2.1 Etapa de Selección de Datos. ....	53
3.2.2 Etapa de Procesamiento de Datos. ....	53
3.2.3 Aplicación Modelo de regresión logística.....	54
3.2.4 Etapa de Sistematización del Modelo Estadístico.....	55
3.2.4.1 Metodología para el desarrollo de la aplicación web. ....	55
<b>4. Resultados.....</b>	<b>57</b>
4.1 Selección de Datos.....	57
4.2 Procesamiento de Datos.....	58
4.2.1 Análisis Univariado.....	58
4.2.1.1 Distribución de Frecuencias. ....	58
4.2.2 Análisis Bivariado.....	72
4.2.2.1 Variables Cuantitativas.....	73
4.2.2.2 Variables Cualitativas.....	74
4.2.2.3 Análisis De Variables Significativas.....	77
4.2.2.4 Perfil del Estudiante desertor.....	84
4.2.3 Análisis Multivariado.....	84
4.2.3.1 Codificación de variables. ....	85
4.2.3.2 Construcción del modelo de regresión logística binaria.....	86

4.2.3.3 Elección del punto de corte. ....	87
4.2.3.4 Aplicación del modelo de regresión logística binaria. ....	90
4.2.3.4 Variables en la ecuación. ....	92
4.2.3.5 Propiedades De Ajuste Y Clasificación Del Modelo. ....	94
4.3 Sistematización Del Modelo Estadístico .....	94
4.3.1 Etapa 1: Formulación. ....	95
4.3.2 Etapa 2: Planificación. ....	95
4.3.2.1 Factibilidad del Sistema. ....	95
4.3.3 Etapa 3: Análisis. ....	97
4.3.4 Etapa 4: Ingeniería. ....	99
4.3.4.1 Diseño de Interfaz.....	100
4.3.4.2 Diseño de Navegación.....	101
4.3.5 Etapa 5: Generación de páginas. ....	101
4.3.6 Etapa 6: Pruebas y Evaluación.....	101
<b>5.Estrategias.....</b>	<b>102</b>
<b>Conclusiones y Recomendaciones .....</b>	<b>104</b>
<b>Bibliografía.....</b>	<b>106</b>
<b>Anexos.....</b>	<b>110</b>

## Lista de Tablas

	Pág.
Tabla 1. Determinantes de la Deserción.....	26
Tabla 2. Frecuencias Observadas.....	29
Tabla 3. Frecuencias Esperadas.....	29
Tabla 4. Concepto OR.....	31
Tabla 5. Operacionalización de Variables.....	49
Tabla 6. Variables obtenidas en la selección de datos.....	57
Tabla 7. Medidas de Tendencia Central y dispersión para variables cuantitativas.....	72
Tabla 8. Asociación de las variables cuantitativas con la variable deserción.....	73
Tabla 9. Asociación de factores individuales con la variable deserción.....	74
Tabla 10. Asociación de factores académicos con la variable deserción.....	75
Tabla 11. Asociación de factores socioeconómicos con la variable deserción.....	75
Tabla 12. Asociación de factores institucionales con la variable deserción.....	77
Tabla 13. Codificación de la variable dependiente.....	85
Tabla 14. Codificación de las variables independientes.....	85
Tabla 15. Recursos Hardware.....	96
Tabla 16. Recursos Software.....	96
Tabla 17. Rol y actividad que cumple el usuario.....	98
Tabla 18. Casos de uso para el módulo del usuario estudiante.....	99

## Lista de Figuras

	Pág.
Figura 1. Concepto Odds Hombres .....	30
Figura 2. Concepto Odds Mujeres.....	31
Figura 3. Distribución por sexo del estudiante.....	58
Figura 4. Distribución por edad de ingreso. ....	59
Figura 5. Distribución por tipo de colegio.....	60
Figura 6. Distribución por puntaje ICFES.....	60
Figura 7. Distribución por promedio de notas.....	61
Figura 8. Distribución por último semestre cursado. ....	61
Figura 9. Distribución por materias perdidas en el último semestre. ....	62
Figura 10. Distribución por total de materias perdida.....	63
Figura 11. Distribución por estrato.....	63
Figura 12. Distribución por lugar de procedencia. ....	64
Figura 13. Distribución por lugar de residencia. ....	65
Figura 14. Distribución por tipo de residencia. ....	65
Figura 15: Distribución por vive con la familia. ....	66
Figura 16. Distribución por hermanos en la universidad. ....	67
Figura 17. Distribución por jefe de familia. ....	67
Figura 18. Distribución por ingresos familiares anuales. ....	68
Figura 19. Distribución por puntaje socioeconómico.....	69
Figura 20. Distribución por valor de matrícula. ....	69
Figura 21. Distribución por tipo de cupo.....	70
Figura 22. Distribución por sanciones.....	71
Figura 23. Deserción estudiantil en el programa de licenciatura en informática .....	71
Figura 24. Porcentaje de deserción según tipo de colegio.....	78
Figura 25. Porcentaje de deserción según estrato social. ....	78
Figura 26. Porcentaje de deserción según tipo de residencia. ....	79
Figura 27. Porcentaje de deserción según vive con la familia. ....	79
Figura 28. Porcentaje de deserción según tiene hermanos en la U. ....	80

Figura 29. Porcentaje de deserción según jefe de familia. ....	80
Figura 30. Porcentaje de deserción según tipo de cupo.....	81
Figura 31. Porcentaje de deserción según sanciones.....	81
Figura 32. Porcentaje de deserción según el valor de matrícula. ....	82
Figura 33. Porcentaje de deserción según promedio de notas.....	82
Figura 34. Porcentaje de deserción según el último semestre cursado.....	83
Figura 35. Porcentaje de deserción según materias perdidas en el último se semestre.....	83
Figura 36. Porcentaje de deserción según el número de materias perdidas en total .....	84
Figura 37. Clasificación inicial del modelo.....	87
Figura 38. Procesamiento de casos.....	88
Figura 39. Curva ROC del modelo.....	88
Figura 40. Área bajo la curva ROC.....	89
Figura 41. Coeficiente K-S.....	90
Figura 42. Resumen del modelo.....	91
Figura 43. Tabla Final de Clasificación. ....	91
Figura 44. Variables de la ecuación. ....	92
Figura 45. Prueba Ómnibus.....	94
Figura 46. Módulo de la aplicación.....	98
Figura 47. Usuario de la aplicación.....	98
Figura 48. Diagrama de casos de uso de usuario Estudiante.....	99
Figura 49. Diseño de Interfaz Gráfica. ....	100
Figura 50. Diseño de Navegación. ....	101

### Lista de Anexos

	Pág.
Anexo 1. Distribución por Género. ....	110
Anexo 2. Distribución por Edad. ....	110
Anexo 3. Distribución por Tipo De colegio. ....	110
Anexo 4. Distribución por Puntaje Icfes. ....	111
Anexo 5. Distribución por Promedio de notas. ....	111
Anexo 6. Distribución por Último semestre cursado. ....	111
Anexo 7. Distribución por Materias perdidas en el último semestre. ....	112
Anexo 8. Distribución por Materias perdidas en total. ....	112
Anexo 9. Distribución por Estrato Social. ....	113
Anexo 10. Distribución por Lugar de procedencia. ....	113
Anexo 11. Distribución por Lugar de residencia. ....	113
Anexo 12. Distribución por Tipo de residencia. ....	114
Anexo 13. Distribución por Vive con la familia. ....	114
Anexo 14. Distribución por Hermanos en la universidad. ....	114
Anexo 15. Distribución por Jefe de familia. ....	115
Anexo 16. Distribución por Ingresos Anuales. ....	115
Anexo 17. Distribución por Puntaje socioeconómico. ....	115
Anexo 18. Distribución por Valor de matricula. ....	116
Anexo 19. Distribución por Tipo de cupo. ....	116
Anexo 20. Distribución por Sanciones. ....	117
Anexo 21. Distribución por deserción estudiantil. ....	117



## Introducción

La deserción estudiantil se ha convertido en un tema de importante relevancia para las Instituciones de Educación Superior por sus efectos dentro del sistema educativo, este problema tiene costos sociales y económicos que afectan a las familias, instituciones educativas y al gobierno en general. En el ámbito personal implica una condición de fracaso que afecta emocionalmente a la persona y sus expectativas profesionales y a nivel institucional comprometen el rendimiento académico de la Universidad.

La incidencia de la deserción universitaria dentro del campo educativo es altamente negativa en la cobertura, puesto que los altos índices de deserción y el aplazamiento de la graduación incurren en la no disponibilidad de cupos para nuevos estudiantes. Un estudiante que abandona sus estudios crea una vacante que posiblemente pudo estar ocupada por otro alumno que se mantuviera estable en su carrera, disminuyendo así la oportunidad de acceso al sistema educativo.

Según el MEN (2012) la tasa de deserción por cohorte, indicador que ilustra la cantidad de estudiantes que desertan de cada 100 que ingresan a algún programa universitario, (cifra sobre la cual se cuenta con mayor información en la región), Colombia alcanza el 45.3%, con valores cercanamente superiores a los presentados por nuestro país están México y Argentina con 42% y 43%, mientras que Venezuela y Chile presentan una deserción del 52% y 54% y Costa Rica con 62%. Esto indica que de cada 100 jóvenes que ingresan, cerca de la mitad terminan desertando. (pág. 2)

La Universidad de Nariño contribuye a la movilidad social de sus estudiantes brindándoles la posibilidad de desarrollar las competencias necesarias para asumir los retos de la sociedad actual. Reconoce que cualquier aspirante puede tener la oportunidad de ingresar a los estudios superiores sin importar su origen o nivel socioeconómico. A pesar de los esfuerzos y políticas públicas para posibilitar el acceso y la permanencia en la universidad, uno de los problemas a los que se ve enfrentada es la deserción estudiantil, por tanto surge un interrogante sobre las causas que conllevan a este fenómeno y cómo se puede determinar el perfil de los estudiantes

desertores. Por ello, es importante realizar estudios que permitan la aplicación de estrategias con el fin de minimizar este fenómeno.

Por esta razón, el presente trabajo tiene como objetivo determinar los factores que inciden en el riesgo de deserción en los estudiantes del programa de Licenciatura en Informática de la Universidad de Nariño y conocer el peso que cada uno de estos tiene a la hora de desertar, para lograrlo se trabaja con un modelo estadístico de clasificación. En primer lugar se encuentran los aspectos generales de la investigación, se hace un recuento sobre lo que ha sido la deserción, resaltando algunas conclusiones importantes de estudios que se han llevado a cabo sobre este aspecto a nivel internacional, nacional y regional. Se continúa con el desarrollo de la metodología que permite la consecución de los objetivos planteados, entre ellos la construcción de un modelo estadístico para calcular el riesgo de deserción en los estudiantes del programa. Por último, se desarrolla una aplicación sencilla con el fin de dar a conocer de forma práctica los resultados obtenidos y se proponen algunas estrategias a partir de este estudio que se podría implementar para reducir los índices de deserción.

## 1. Aspectos Generales

### 1.1 Problema de Investigación

#### 1.1.1 Planteamiento del Problema.

La culminación de los estudios superiores de un estudiante es el cumplimiento de los objetivos de una Institución de Educación Superior, lo cual se logra desarrollando procesos formativos e implementación de estrategias de apoyo en todos los ciclos de su permanencia, pero esto se puede ver afectado por la deserción estudiantil, uno de los principales problemas del sector educativo.

Según Tinto & Giovagnoli referenciados por el MEN (2009) se puede entender la deserción como una situación a la que se enfrenta un estudiante cuando aspira y no logra concluir su proyecto educativo, considerándose como desertor a aquel individuo que siendo estudiante de una Institución de Educación Superior no presenta actividad académica durante dos semestres consecutivos, lo cual equivale a un año de inactividad académica. (pág. 22)

Por mucho tiempo la deserción universitaria se consideró como algo habitual, incluso representaba la exigencia y calidad de una institución, en la actualidad se puede afirmar que es todo lo contrario ya que representa una falta de compromiso y generación de políticas hacia los estudiantes, además de producir grandes costos, lo cual se ha convertido en un problema que se necesita combatir.

La deserción es una problemática que aflige a la Universidad de Nariño y en particular al programa de Licenciatura en Informática, ya que la cantidad de estudiantes que ingresan al programa no es la misma que culmina satisfactoriamente sus estudios superiores. Esto se debe a la escasa intervención y seguimiento apropiado del fenómeno, así como su estudio desde diferentes dimensiones, lo que permitiría identificar con mayor precisión los factores determinantes que inciden en el riesgo de deserción. Según el MEN (2016) en su Boletín Educación Nacional en cifras: “La identificación de estos factores permite a las instituciones de educación superior contar con elementos fundamentales que orienten sus planes de acción, de esta forma implementar estrategias integrales que fomenten la permanencia desde sus realidades y necesidades” (pág. 3).

En el programa de Licenciatura en informática es necesario ampliar la comprensión del problema con un enfoque diferente, para conocer la información que permita encontrar los factores que llevan a un estudiante a desertar y con ello poder desarrollar estrategias de retención y permanencia para evitar el abandono de estudios y mejorar la calidad educativa.

### **1.1.2 Formulación del Problema.**

¿Cuáles son los factores que inciden en el riesgo de deserción en los estudiantes del programa de Licenciatura en Informática de la Universidad de Nariño?

## **1.2 Objetivos**

### **1.2.1 Objetivo General.**

Identificar los factores asociados a la deserción estudiantil en el programa de Licenciatura en Informática de la Universidad de Nariño, para construir y sistematizar un modelo de regresión logística a partir del cual se pueda plantear estrategias de retención.

### **1.2.2 Objetivos Específicos.**

- Identificar los factores más importantes que inciden en la deserción y establecer su grado de influencia.
- Construir un modelo de regresión logística que permita calcular la probabilidad de deserción en los estudiantes del programa de Licenciatura en Informática.
- Sistematizar el modelo obtenido por medio de una aplicación que permita estimar automáticamente el riesgo de deserción de los estudiantes.
- Proponer estrategias para reducir el problema de la deserción en la población estudiantil del programa.

### 1.3 Justificación

Uno de los desafíos en materia de educación universitaria es garantizar el acceso y la permanencia de los jóvenes, sin distinciones de etnia, género, ideología, religión, condición socioeconómica u otras circunstancias, todos tienen la oportunidad de participar y tener éxito en la educación superior, contribuyendo así en la construcción de una sociedad justa y equitativa. La deserción universitaria se ha convertido en un dolor de cabeza para las instituciones, quienes deben enfrentar el problema generando soluciones que permitan contrarrestar sus efectos.

Las investigaciones asociadas a los factores de deserción universitaria realizadas por el Ministerio de Educación Nacional y el Gobierno, partiendo de una revisión exhaustiva de la literatura existente han arrojado importantes estudios que permitieron la construcción del estado de arte sobre la deserción estudiantil, el cual demuestra la pluralidad de las causas de este problema por las cuales los estudiantes se ven obligados a abandonar sus carreras impidiendo así su graduación. La falta de profundización sobre estos factores, influye en la poca prevención hacia los estudiantes con alto riesgo de deserción.

La revisión de las cifras existentes sobre la deserción estudiantil colombiana lleva a la conclusión de que aún es un problema dentro de las instituciones de educación superior, además de no contar con la respuesta adecuada a este fenómeno. Faltan herramientas efectivas para cuantificar los riesgos y generar estrategias oportunas para minimizarlos, de modo que los estudiantes logren permanecer en la institución y graduarse de los programas inscritos.

El problema de la deserción en la Universidad de Nariño y específicamente en el programa de Licenciatura en Informática merece un estudio minucioso, para que a partir de la información que se disponga sobre los desertores se pueda emprender actividades válidas que permitan la permanencia de los estudiantes.

La oficina de Registro y Control Académico de la Universidad de Nariño (OCARA) cuenta con información detallada de cada estudiante, la cual es muy útil para determinar los factores más relevantes que inciden en la deserción, es necesario que se realice una investigación para

determinar cuál es perfil de un estudiante desertor y las causas asociadas a este fenómeno, con el fin de generar estrategias orientadas a disminuir las cifras de deserción, mejorando así la calidad de los programas académicos. La importancia de este trabajo, radica en la sistematización de un instrumento basado en el modelo de regresión logística que a partir de los factores incidentes, permita determinar el riesgo de deserción en los estudiantes del programa, contribuyendo así a la generación de las estrategias para disminuirlo.

## 2. Marco Referencial

### 2.1 Marco Teórico

#### 2.1.1 Deserción Estudiantil Universitaria.

El tema de deserción estudiantil universitaria es complejo y no existe una única definición de deserción que capte la totalidad de su dimensión. Por esto, se intenta determinar desde diferentes perspectivas y causas que llevan al abandono de la educación superior. Tinto (1989) afirma:

El estudio de la deserción de la educación superior es extremadamente complejo, pues implica no solo una variedad de perspectivas sino también una gama de diferentes tipos de abandono. Probablemente ninguna definición puede captar en su totalidad la complejidad de este fenómeno universitario. Los investigadores y funcionarios de instituciones deben elegir con cuidado las definiciones que mejor se ajusten a sus intereses y metas. (pág. 9)

Estas perspectivas dependen de las partes interesadas, entonces, se define la deserción como un comportamiento individual, desde el punto de vista institucional y por último una perspectiva estatal o nacional.

El punto de partida para desarrollar una definición de deserción adecuada a la perspectiva del individuo, es el conocimiento de que los significados que un estudiante asigna a su comportamiento pueden diferir sustancialmente de los que un observador atribuye a ese mismo comportamiento. (Tinto, 1989, pág. 1)

De acuerdo con ello, desertar según un estudiante puede ser un paso positivo hacia el logro de una meta, mientras que para los funcionarios puede definirse como el fracaso en complementar el programa de estudios. Por tanto el punto de vista individual difiere a la posición del otro individuo, de acuerdo con los intereses, propósitos y metas que tiene al incorporarse al sistema educativo. El término de deserción en este contexto también puede ser modificado durante el transcurso de la carrera dependiendo de otros factores como una mayor madurez individual o por la experiencia adquirida.

Para Tinto (1989) Desde el punto de vista individual, desertar significa el fracaso para completar un determinado curso de acción o alcanzar una meta deseada, en pos de la cual el sujeto ingresó a una particular institución de educación superior. Por consiguiente, la deserción no sólo depende de las intenciones individuales sino también de los procesos sociales e intelectuales a través de los cuales las personas elaboran metas deseadas en una cierta universidad. (pág. 3)

Desde la perspectiva institucional es más sencillo definir deserción que desde un punto de vista individual. “Todos los sujetos que abandonan una institución de educación superior pueden, teniendo en cuenta las razones alegadas para hacerlo, ser clasificados como desertores” (Tinto, 1989, pág. 5). Este concepto permite identificar a todas las personas que abandonan una institución antes de culminar sus estudios y no recibir su título universitario. Cuando se refiere a la deserción desde la perspectiva estatal o nacional cambia notablemente porque depende de la estructura de un país.

Para efectos de este estudio se considera la deserción como el abandono o retiro ya sea voluntario e involuntario de un estudiante, implica transferencias a otros programas o instituciones universitarias. La deserción voluntaria se da cuando el estudiante renuncia al programa o lo abandona sin dar información a la universidad, mientras que la no voluntaria se presenta cuando la institución decide según el reglamento y por algún motivo ya sea disciplinario o académico desvincular al estudiante del programa.

### **2.1.2 Tipos de Deserción Estudiantil.**

El fenómeno de la deserción estudiantil se encuentra en cualquier institución educativa, independiente de las características que estas tengan, a través de la exploración de la temática de la deserción y según el Ministerio de Educación Nacional podemos reconocer dos tipos de abandono: uno con respecto al tiempo y otro con respecto al espacio.

La deserción con respecto al tiempo según el MEN (2009) se clasifica en:

#### **2.1.2.1 Deserción precoz.**

“Individuo que habiendo sido admitido por la institución de educación superior no se matricula” (pág. 22). Cuando un estudiante tiene un primer contacto con la universidad se



forman las primeras impresiones sobre las características de ésta, los factores como la desinformación y la veracidad de la misma pueden influir para que ocurra este suceso.

#### ***2.1.2.2 Deserción temprana.***

“Individuo que abandona sus estudios en los primeros semestres del programa” (pág. 22). Este tipo de deserción se da en los primeros años de la duración reglamentaria del programa, ocurre cuando un estudiante está en proceso de cambio de colegio a la universidad, aquí hay una adaptación social y académica a la vida universitaria. Es posible que se genere en el estudiante expectativas equivocadas de lo que ofrece el programa o existan razones diferentes en la que la institución no puede intervenir.

#### ***2.1.2.3 Deserción tardía.***

“Individuo que abandona los estudios en los últimos semestres” (pág. 22). Se presenta esta deserción entre los últimos años de la carrera universitaria y antes de la culminación de la misma, existen casos donde los estudiantes presentan problemas y dificultades que no son capaces de superar. Aunque son abundantes los factores determinantes en la deserción, en este periodo hay una disminución sustancial de ésta, porque pesa mucho más el tiempo que el estudiante ha invertido para llegar hasta este punto.

Por su parte, según el MEN (2009, pág. 24) la deserción con respecto al espacio se divide en:

#### ***2.1.2.4 Deserción institucional.***

“Caso en el cual el estudiante abandona la institución”.

#### ***2.1.2.5 Deserción interna o del programa académico.***

“Se refiere al alumno que decide cambiarse a otro programa que ofrece la misma institución de educación superior”. Donde los factores internos y externos al estudiante determinan el éxito o el fracaso en la culminación de los estudios superiores.

### 2.1.3 Factores de Deserción.

Para estudiar el fenómeno de deserción es necesario identificar los factores que la explican, según en la temática encontrada sobre el tema de estudio, se entiende que los factores son diversos debido a la multiplicidad de estudiantes dentro de una institución, también dependen del momento en el que se presentan al programa.

Sin embargo, se puede decir que la deserción es el resultado de un conjunto de factores, desde diferentes perspectivas en la que se destacan las variables institucionales, socioeconómicas, académicas y personales las cuales se examinan individualmente.

**Tabla 1**

*Determinantes de la Deserción*

<b>Individuales</b>	<b>Socioeconómicos</b>	<b>Académicos</b>	<b>Institucionales</b>
·Edad, género y estado civil.	·Estrato social	·Orientación profesional	·Normatividad académica
·Calamidad y/o problema doméstico.	·Situación laboral del estudiante	·Tipo de colegio de secundaria	·Becas y formas de financiamiento
·Integración social	·Situación laboral de los padres	·Rendimiento académico superior	·Recursos universitarios
·Expectativas no satisfechas	·Dependencia económica	·Métodos de estudio	·Relación con el profesorado y con los demás estudiantes
·Incompatibilidad horaria con actividades extra académicas	·Personas a cargo	·Calificación en el examen de admisión	·Grado de compromiso con la institución educativa
	·Nivel educativo de los padres	·Insatisfacción con el programa académico	·Calidad de programa
	·Entorno familiar	·Carga académica	
	·Entorno macroeconómico del país	·Repitencia	

Fuente: (Vásquez, Castaño, Gallón, & Gómez, 2003)

La anterior tabla nos resume los diferentes estudios y perspectivas de la deserción, donde están las variables más determinantes divididas en cuatro categorías precisas del problema, las cuales serán consideradas para la presente investigación, sin embargo las variables a examinar dentro de cada grupo de factores dependen de los datos disponibles en las fuentes de información.

#### **2.1.4 Análisis Estadístico.**

“Uno de los primeros procedimientos estadísticos que el investigador o administrador debe realizar es la descripción de los datos y la identificación de patrones básicos de los mismos” (Castañeda, Cabrera, Navarro, & Vries, 2010, pág. 26) Los datos son el pilar fundamental para el análisis estadístico, por ende es sumamente importante hacer una descripción exhaustiva de los mismos, para ello se tiene en cuenta una serie de técnicas estadísticas que permiten interpretarlos y descubrir patrones de comportamiento. El análisis estadístico puede ayudar a explicar las circunstancias que determinan la ocurrencia de algún fenómeno y se divide en tres grandes grupos:

##### ***2.1.4.1 Análisis Univariado.***

También conocido como estadística descriptiva, se concentra en cada una de las variables de modo independiente, de acuerdo con el tipo de variable se usa distribuciones de frecuencias, medidas de tendencia central y de dispersión.

##### ***2.1.4.2 Análisis Bivariado.***

Se utiliza para determinar la existencia o ausencia de asociación entre dos variables, permite examinar el efecto o influencia de una variable independiente sobre una dependiente, enfatizando principalmente en la forma como se distribuye la segunda en función de las categorías de la primera.

El grado de relación entre las variables categóricas no se puede establecer simplemente con la observación de las frecuencias, para ello se debe utilizar alguna medida de asociación. Uno de los procedimientos estadísticos diseñados para este propósito es la prueba Chi cuadrado de Pearson.

##### ***2.1.4.3 Análisis Multivariado.***

Dentro de las técnicas multivariantes encontramos los métodos de dependencia los cuales adoptan dos grupos de variables; dependientes e independientes. Estos procedimientos permiten analizar simultáneamente los datos para determinar el grado de influencia de dos o más variables independientes sobre una variable dependiente. Uno de los métodos más utilizados se denomina; regresión logística.

### 2.1.5 Chi- cuadrado de Pearson.

Es una prueba de significación estadística que proporciona una expresión numérica conocida como  $X^2$ , propuesta por Pearson en 1911, este valor indica la relación existente entre dos variables ya sean politómicas o dicotómicas y se calcula mediante la presentación de datos en tablas de contingencia. Para Barón & Téllez (2004) “la prueba de independencia ji-cuadrado (chi-cuadrado) contrasta la hipótesis de que las variables son independientes, frente a la hipótesis alternativa de que una variable se distribuye de modo diferente para diversos niveles de la otra” (pág. 44). Según Pardo & Ruiz (2005) “para ello, compara las *frecuencias observadas* (frecuencias de hecho obtenidas) con las *frecuencias esperadas* (las frecuencias que teóricamente deberíamos haber encontrado en cada casilla si los dos criterios de clasificación fueran independientes)” (pág. 12).

Cuando las variables son independientes las frecuencias esperadas se estiman de la siguiente forma:

$$(Frecuencia\ esperada)_{ij} = \frac{(Total\ de\ la\ fila\ i) \times (Total\ de\ la\ columna\ j)}{No.\ total\ de\ casos}$$

Donde  $i$  hace referencia a una fila cualquiera,  $j$  a una columna cualquiera,  $ij$  a una casilla cualquiera. Es decir, la frecuencia esperada de una casilla se obtiene dividiendo el producto de total fila y total columna correspondientes a esa casilla sobre el número total de casos.

Una vez obtenidas las frecuencias esperadas en cada casilla, el estadístico  $X^2$  o *Chi-cuadrado* de Pearson se obtiene de la siguiente manera:

$$X^2 = \sum_i \sum_j \frac{(n_{ij} - m_{ij})^2}{m_{ij}}$$

Donde  $n_{ij}$  hace referencia a las frecuencias observadas y  $m_{ij}$  a las esperadas. (Pardo & Ruiz, 2005, pág. 212)

El valor obtenido representa la magnitud del error en caso de que aceptemos la hipótesis alterna (existe relación entre las variables), entre más alto sea el valor  $X^2$  (mayor a 0,05) más

posibilidad de independencia, por el contrario entre más próximo sea  $X^2$  a cero (menor a 0,05) más probabilidad de que las variables están relacionadas.

La siguiente tabla nos muestra información sobre dos variables y sus respectivas frecuencias observadas.

**Tabla 2**

***Frecuencias Observadas***

<i>Frecuencias Observadas</i>		<i>Evento</i>		<i>Total</i>
		<i>Si</i>	<i>No</i>	
Variable A	Categoría 1	$a_o$	$b_o$	$a_o + b_o = x$
	Categoría 2	$c_o$	$d_o$	$c_o + d_o = y$
<i>Total</i>		$a_o + c_o = i$	$b_o + d_o = j$	$x + y = n$

Fuente: Esta investigación

Aplicando la fórmula para calcular las frecuencias esperadas se obtiene la siguiente información:

**Tabla 3**

***Frecuencias Esperadas***

<i>Frecuencias Esperadas</i>		<i>Evento</i>		<i>Total</i>
		<i>Si</i>	<i>No</i>	
Variable A	Categoría 1	$\frac{(x * i)}{n} = a_e$	$\frac{(x * j)}{n} = b_e$	$a_e + b_e$
	Categoría 2	$\frac{(y * i)}{n} = c_e$	$\frac{(y * j)}{n} = d_e$	$c_e + d_e$
<i>Total</i>		$a_e + c_e$	$b_e + d_e$	$(a_e + b_e) + (c_e + d_e)$

Fuente: Esta investigación

Luego se aplica la fórmula del estadístico  $X^2$  a cada celda de la tabla, en el caso de que ocurra el evento en categoría 1, el contenido es:

$$X^2 = \frac{(a_o - a_e)^2}{a_e}$$

Este procedimiento se hace para cada celda y luego se suman los valores obtenidos, encontrando así el estadístico  $X^2$  de Pearson.

Para que las probabilidades de distribución  $\chi^2$  constituyan una buena aproximación a la distribución del estadístico  $X^2$  conviene que se cumplan algunas condiciones; entre ellas, que las frecuencias esperadas no sean demasiado pequeñas. Suele asumirse que, si existen frecuencias esperadas menores que 5 estas no deben superar el 20% del total de frecuencias esperadas. (Pardo & Ruiz, 2005, pág. 212)

El estadístico  $X^2$  indica si existe o no relación entre dos variables pero no la magnitud ni la dirección de esta asociación, para ello se utiliza otros estadísticos como el índice de riesgo.

### 2.1.6 Odds Ratio.

La estrategia de recogida de datos del presente estudio, está relacionada con un diseño longitudinal *hacia atrás*, también llamado estudio retrospectivo o de caso-control, en la cual se forman dos grupos; desertores y no desertores. A partir de esta información se hace un seguimiento para encontrar los factores desencadenantes. Como el tamaño de los grupos se fija según la presencia o ausencia de la deserción se calcula el OR-*odds ratio* (razón de las ventajas o razón de productos cruzados) el cual se define como: “un cociente entre dos odds” (Cerdeira, Vera, & Rada, 2013, pág. 1330). Para Borracci (2001) “El uso preferencial del OR se debe a que éste proporciona un enlace entre los estudios de cohortes y los de caso control, y a su vez con el modelo de regresión logística, principal método de análisis para los últimos” (pág. 667).

El concepto de *odds* es una forma de expresar la posibilidad de ocurrencia de un evento. A modo de ejemplo: en una muestra de 100 estudiantes de género masculino, 75 de ellos han culminado sus estudios exitosamente, la probabilidad de graduarse sería  $P = 75/100 = 0,75$  y la de desertar  $P = 25/100 = 0,25$ . Al dividir estas dos probabilidades se obtiene la *Odds* de graduación, que en este caso tendría un valor de 3, es decir, por cada estudiante de este género que abandonó sus estudios 3 lograron graduarse.

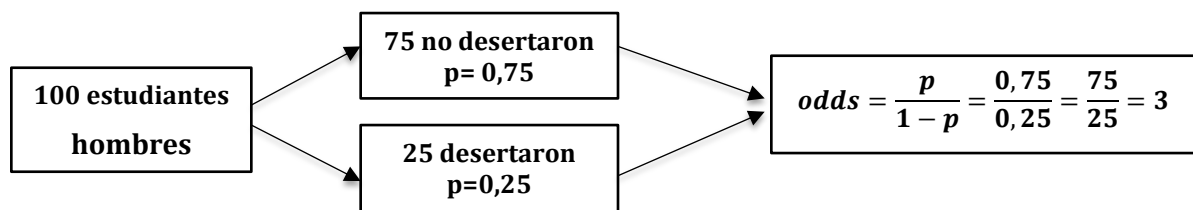


Figura 1. Concepto Odds Hombres  
Fuente: Esta investigación

El *Odds ratio* es la división entre 2 *Odds*, en el ejemplo anterior de 100 estudiantes hombres 75 lograron graduarse ( $Odds = 75/25 = 3$ ). Ahora supongamos que hay 85 estudiantes de género femenino de las cuales 68 lograron graduarse, la Odds esta vez sería de 4 ( $Odds = 68/17 = 4$ ), como lo muestra la figura 2.

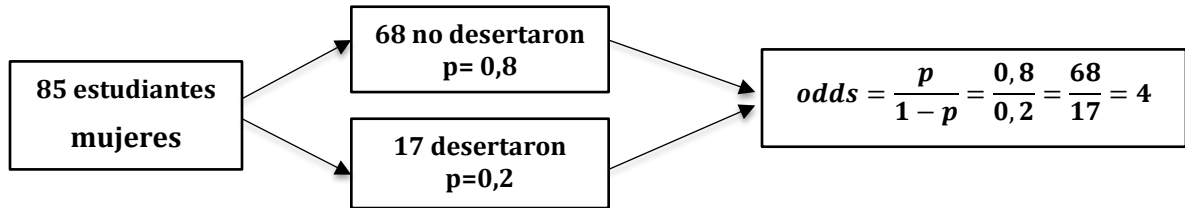


Figura 2. Concepto Odds Mujeres  
Fuente: Esta investigación

La OR (**Odds Ratio**) se obtiene al dividir la Odds de los hombres por la Odds de las mujeres:

$$OR = \frac{Odds\ Mujeres}{Odds\ Hombres} = \frac{4}{3} = 1,3$$

En este caso la OR equivale a 1,3 para las mujeres respecto a los hombres, por tanto la OR es el resultado de la división entre dos Odds.

Basándose en la siguiente tabla de frecuencias se puede indicar otra forma de expresar la OR la cual sería:

Tabla 4

**Concepto OR**

		Graduación		Total
		Si	No	
Sexo	Femenino	68	17	85
	Masculino	75	25	100
Total		143	42	185

Fuente: Esta investigación

$$OR = \frac{Odds\ Mujeres}{Odds\ Hombres} = \frac{(68 * 25)}{(17 * 75)} = \frac{1700}{1275} = 1,3$$

Es importante resaltar que la odds de referencia debe ubicarse en el denominador y la del grupo de interés en el numerador. Una OR=1 indica que las dos categorías son iguales; es un valor nulo. Cuando la OR es menor a 1 indica que el evento es menos frecuente en la categoría de interés si en la de referencia ocurriera por azar.

### **2.1.7 Regresión Logística.**

Para Alderete (2006) “La regresión logística, al igual que otras técnicas estadísticas multivariadas, da la posibilidad de evaluar la influencia de cada una de las variables independientes sobre la variable dependiente o de respuesta y controlar el efecto del resto” (pág. 53). El objetivo principal del modelo de regresión logística es determinar el grado de influencia de los diferentes factores en la probabilidad de un suceso, este modelo es útil cuando la variable de respuesta es dicotómica, es decir, toma dos valores; presencia o ausencia, las variables independientes pueden ser cuantitativas o cualitativas.

El análisis de regresión logística utiliza el mismo método que la regresión lineal múltiple, la diferencia es que en esta última la variable dependiente es métrica, aunque ambas técnicas son muy parecidas los enfoques son diferentes.

### **2.1.8 Regresión logística binaria.**

Cuando la variable dependiente es dicotómica como en el caso del presente estudio se realiza un proceso binomial, ya que se tiene dos posibles resultados, si desertó o no desertó. El modelo determina una probabilidad la cual va a pertenecer a uno de los dos grupos establecidos, este diagnóstico se define de acuerdo con una serie de variables predictoras o independientes de cada estudiante. Una de las ventajas de la regresión logística binaria es que no se necesita plantear una hipótesis de partida, así como tampoco es necesario comprobar la normalidad en la distribución de las variables lo cual mejora su tratamiento.

Para Barón & Téllez (2004) el modelo de regresión logística puede escribirse como:

$$\text{Log} \left( \frac{p}{1-p} \right) = b_0 + b_1 x_1 + \dots + b_n x_n$$



Donde  $p$  es la probabilidad (riesgo) de que ocurra el evento de interés, en este caso la deserción, las variables independientes están representadas con la letra  $x$ , y los coeficientes asociados a cada variable con la letra  $b$ . (pág. 54)

Dado el valor de las variables independientes se puede calcular directamente la estimación del riesgo de deserción de un sujeto  $i$  de la siguiente forma:

$$P_i = \frac{e^z}{1+e^z} \quad \text{Que es igual a} \quad P_i = \frac{1}{1+e^{-z}} \quad \text{Por lo que para el sujeto } i$$

$$P_i(Y = 1) = \frac{e^{(\beta_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_n x_n)}}{1 + e^{(\beta_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_n x_n)}}$$

Donde  $\beta_0, b_1, b_2, b_3, b_n$  son los parámetros del modelo, y  $e$  denota la función exponencial. Esta función exponencial es una expresión simplificada que corresponde a elevar el número  $e$  a la potencia contenida dentro del paréntesis, siendo  $e$  el número o constante de Euler, o base de los logaritmos neperianos cuyo valor aproximado a la milésima es 2,718. (Aguayo, 2007, pág. 2)

### 2.1.9 Curva ROC.

Para Del Valle La curva ROC es una herramienta estadística utilizada en el análisis de la clasificación de la capacidad discriminante de una prueba diagnóstica dicotómica. Es decir, una prueba, basada en una variable de decisión, cuyo objetivo es clasificar a los individuos de una población en dos grupos: uno que presente un evento de interés y otro que no. Esta capacidad discriminante está sujeta al valor umbral elegido de entre todos los posibles resultados de la variable de decisión, es decir, la variable por cuyo resultado se clasifica a cada individuo en un grupo u otro. (pág. 6)

En el gráfico de la curva ROC se observan los pares de Sensibilidad () y Especificidad () resultantes de la variación de los puntos de corte en todo el rango de los resultados observados. En el eje ( $y$ ) de coordenadas se sitúa la sensibilidad o fracción de verdaderos positivos, En el eje ( $x$ ) se sitúa la fracción de falsos positivos o 1- especificidad. Cuanto más alejada del eje de abscisas esté la curva que se genera uniendo estos puntos, más eficiente resulta la prueba

diagnóstica. El área bajo la curva ROC indica la probabilidad de que se distinga correctamente a un individuo en la ocurrencia o no del evento, es por ello que el peor resultado es el que tiene una curva ROC que coincide con la diagonal, por debajo de la cual hay un área de 0,5.

### **2.1.10 Estrategias.**

El creciente acceso de estudiantes más pobres a la educación superior implica mayor vulnerabilidad y exposición a la deserción. Por ende, su tratamiento también debe tener una mayor relevancia en las Instituciones de Educación Superior. A lo anterior se suman las metas fijadas tanto en el Plan Decenal de Educación como en el documento Visión Colombia II Centenario 2019, de pasar de una tasa actual de deserción del 48.4% en 2004 a una del 25% en 2019 (MEN, 2009).

La principal unidad de análisis debe ser el estudiante y esta información es indispensable para conocer su condición y los factores que pueden afectar a lo largo de su carrera universitaria, teniendo en cuenta las causales de deserción se generan estrategias para contrarrestar ese fenómeno. Estas estrategias de permanencia estudiantil, dependen de los factores asociados a la eficiencia educativa que promueven a un estudiante terminar sus estudios, de acuerdo con esto, son diferentes circunstancias internas y externas a ese individuo que influyen en su permanencia.

Este importante tema está contemplado con diferentes proyectos y estrategias por el Ministerio de Educación Nacional y las diferentes instituciones, para concentrar esfuerzos en el acceso y permanencia a la educación superior con resultados que traduzcan a la graduación de los estudiantes.

En cuanto a las instituciones educativas, debido a la pluralidad de sus estudiantes y de la variabilidad de los factores que componen la deserción, las estrategias que se ejecutan son integrales y corresponden a la caracterización individualizada, entre ellas están el bienestar universitario con actividades para el desarrollo de los estudiantes, programas de acompañamiento y apoyo, gestión académica y universitaria entre otras.

## 2.2 Marco Conceptual

**Deserción Estudiantil.** “La deserción puede definirse como el abandono definitivo de las aulas y del programa de formación” (MEN, 2015, pág. 15). Para efectos de estudio se entiende como el retiro voluntario o forzado de la carrera en la cual un estudiante se encuentra matriculado, sin terminar las actividades académicas. Este abandono puede presentarse de diversas maneras ya sea por decisión personal o porque la universidad según su estatuto se ve obligada a retirar al estudiante. Por lo tanto, el índice de deserción se refiere a la cantidad de estudiantes que no culminaron un semestre académico o que no se matricularon en el siguiente.

**Desertor.** “Desde el punto de vista institucional todos los estudiantes que abandonan una institución de educación superior pueden ser clasificados como desertores; en este sentido, muchos autores asocian la deserción con los fenómenos de bajo rendimiento académico y retiro forzoso” (MEN, 2009, pág. 20). El desertor es una persona que posterga sus metas académicas debido a diferentes factores ya sean intrínsecos o extrínsecos, por los cuales se ve obligado a abandonar la institución antes de culminar sus estudios.

**Nivel de deserción.** Es el semestre académico en el cual el estudiante abandona sus estudios ya sea de manera voluntaria o por retiro forzoso, para los estudiantes que en el momento de la deserción están cursando materias de diferentes semestres el nivel será estimado de acuerdo con el semestre en el que se encuentra matriculado.

**Retención.** “Habilidad de una institución para que el estudiante permanezca en ella desde la admisión hasta la graduación” (MEN, 2015, pág. 15). La Institución Universitaria debe buscar acciones que conduzcan a la reducción de la deserción estudiantil, estableciendo estrategias que garanticen el bienestar del alumno, brindando acompañamiento académico, apoyo socioeconómico, psicológico, atención en salud, recreación y deporte que permitan la permanencia y finalización de los estudios en los diferentes programas.

**Bienestar Estudiantil.** “Se asocia con la presencia de factores protectores que elevarían la probabilidad de evitar daños específicos en una población, producto de la manifestación de una condición vulnerable” (Castro & Rivas, 2006, pág. 39). El direccionamiento de Bienestar Estudiantil, está orientado a posibilitar una mejor calidad de vida y el desarrollo integral de los

estudiantes, donde se reconozca como condición básica el ser humano, y permita construir espacios que afiancen las actividades físicas, mentales, espirituales y sociales.

La Universidad de Nariño cuenta con el sistema de bienestar universitario, el cual es el encargado de orientar y brindar apoyo hacia el mejoramiento de las condiciones de vida de la comunidad universitaria, mediante el fortalecimiento de las capacidades humanas y su proyección social.

**Factores de riesgo.** “Causas que inciden en el proceso de forma tal que contribuyen a que la probabilidad de abandono del sistema escolar sea mayor” (Castro & Rivas, 2006, pág. 41). Los factores asociados a la deserción escolar están relacionados con las condiciones y situaciones de vida que experimentan los jóvenes, los cuales fortalecen o disminuyen la probabilidad de tomar la decisión de abandonar los estudios.

**Software SPSS.** El SPSS (*Statistical Product and Service Solutions*) es una potente herramienta de tratamiento de datos y análisis estadístico. Al igual que el resto de aplicaciones que utiliza como soporte el sistema operativo Windows, el SPSS funciona mediante menús desplegables y cuadros que permiten hacer la mayor parte del trabajo utilizando el puntero del ratón. (Pardo & Ruiz, 2005, pág. 3)

**Variable cuantitativa.** “Son las variables que pueden medirse, cuantificarse o expresarse numéricamente” (Pita & Pértega, 1997, pág. 2).

**Variable cualitativa.** Este tipo de variables representan una cualidad o atributo que clasifica a cada caso en una de varias categorías. La situación más sencilla es aquella en la que se clasifica cada caso en uno de dos grupos (hombre/mujer, enfermo/sano, fumador/no fumador). Son datos dicotómicos o binarios. (Pita & Pértega, 1997, pág. 2)

Para De la Fuente (2011) las variables cualitativas se pueden clasificar de varias formas:

### ***1. Según el número de categorías***

**Dicotómicas:** Solo hay dos modalidades. Ejemplo, padecer una enfermedad (Sí, No), Sexo (Hombre, Mujer), en general los fenómenos de respuesta binaria.

*Politómicas:* Cuando hay más de dos categorías. Ejemplo, fenómenos de respuesta múltiple, lugar de nacimiento, clase social, etc.

## **2. Según la escala de medida de las categorías**

*Nominal:* No se puede definir un orden natural entre sus categorías, por ejemplo, la raza (blanca, negra, otra), etc.

*Ordinal:* Es posible establecer relaciones de orden entre las categorías. Por ejemplo, la clase social (alta, media, baja), etc.

*Por Intervalo:* Proceden de variables cuantitativas agrupadas en intervalos, ejemplos de este tipo son el sueldo, la edad, los días del mes, etc.

**Variable Independiente.** Es una variable que puede tener su origen en el sujeto o en el entorno del sujeto. Es la variable que el investigador manipula para ver los efectos que produce en otra variable. En la relación más simple el investigador estudia qué le sucedería a la variable efecto cuando cambian los valores de la variable causa o variable independiente. (Buendía, Cólás, & Hernández, 2001, pág. 3)

**Variable dependiente.** “Es el factor que el investigador observa o mide para determinar el efecto de la variable independiente. La variable dependiente es la variable de respuesta o salida” (Buendía, Cólás, & Hernández, 2001, pág. 3)

**Variable dummy.** El modelo de regresión logística utiliza variables cuantitativas y cualitativas, estas últimas se miden a través de datos categóricos, para poderlos incorporar en el modelo de regresión se crea variables ficticias de tipo dicotómico, llamadas también “*variables dummy*”. Para Murillo & González (2000) las variables ficticias traducen esos factores cualitativos en variables exógenas en la ecuación de regresión. El aspecto más destacable de este tipo de variables es que se trata de variables expresadas, en cualquier caso, en forma de variables binarias que toman en modelo econométrico los valores 0 ó 1. Este código binario es asumido independientemente del número de modalidades, o categorías, del factor cualitativo representado. (pág. 3)

**Tablas de contingencia.** Cuando se trabaja con variables categóricas los datos suelen organizarse en tablas de doble entrada en las que cada entrada representa un criterio de clasificación (una variable categórica). Como resultado de esta clasificación las frecuencias (el número o porcentaje de casos) aparecen organizadas en casillas que contienen información sobre la relación existente entre ambos criterios. (Pardo & Ruiz, 2005, pág. 207)

Las tablas de contingencia son una de las formas más comunes de resumir los datos categóricos, su finalidad es analizar si existe asociación entre las variables y calcular la intensidad de esta asociación.

**Independencia.** La independencia de dos variables consiste en que la distribución de una de las variables es similar sea cual sea el nivel que examinemos de la otra. Esto se traduce en una tabla de contingencia en que las frecuencias de las filas (y las columnas) son aproximadamente proporcionales. (Barón & Téllez, 2004, pág. 44)

**Modelo de Regresión Logística.** “El modelo de regresión logística permite estimar la probabilidad de un suceso que depende de los valores de ciertas covariables” (Cuadras, 2014, pág. 233). En estadística es un tipo de análisis de regresión utilizado para predecir el resultado de una variable categórica (aquella variable que pueda adoptar un número ilimitado de categorías) en función de variables predictoras.

**Pruebas de Bondad de Ajuste.** Según Berrocal (2006) “La bondad de ajuste es un indicador que permite discernir acerca de que tan buena es la ecuación obtenida. Para determinar la bondad de ajuste se utilizan diferentes criterios de la regresión lineal” (pág. 7). Algunos indicadores de esta bondad de ajuste son realizados mediante la prueba de ji-cuadrado ( $X^2$ ) y Kolmogorov-Smirnov (K-S) entre otros.

**Sistematización.** Para Jara citado por Granados (2005) La sistematización es aquella interpretación crítica de una o varias experiencias, que a partir de su ordenamiento y reconstrucción, descubre o explicita la lógica del proceso vivido, los factores que han intervenido en dicho proceso, cómo se han relacionado entre sí, y por qué lo han hecho de ese modo. (pág. 25)

La sistematización busca dar a conocer la realidad de lo encontrado a largo de la investigación con el fin de que ésta trascienda, para esto es necesario lograr un equilibrio entre la teoría y la práctica, que requiere de un proceso dinámico, creativo, riguroso y preciso, para diseñar una propuesta que pueda cubrir la necesidad de organización de la información de manera que se torne útil y contribuya a la solución de un problema.

### **2.3 Marco Legal**

La ley 115 (1994) establece en su Artículo 1o. “La educación es un proceso de formación permanente, personal, cultural y social que se fundamenta en una concepción integral de la persona humana, de su dignidad, de sus derechos y de sus deberes.” Señala las normas generales para regular el Servicio Público de la Educación que cumple una función social acorde con las necesidades e intereses de las personas, de la familia y de la sociedad. Esta ley reconoce los derechos de los estudiantes en la parte humana e integral que pueden estar relacionadas a los factores que inciden en la deserción y que influyen en la determinación de no continuar con la educación.

El marco legal de educación superior Colombiana se transformó a partir de la entrada en vigencia de la ley 30 de (1992) que dio autoridad a las universidades para su constitución y generación de programas académicos. En el Artículo 1° plantea: La Educación Superior es un proceso permanente que posibilita el desarrollo de las potencialidades del ser humano de una manera integral, se realiza con posterioridad a la educación media o secundaria y tiene por objeto el pleno desarrollo de los alumnos y su formación académica o profesional. Es así como el principal objetivo de la educación superior es lograr que sus estudiantes culminen con satisfacción sus carreras académicas.

El estado y la Constitución con la presente ley garantizan la autonomía universitaria y vela por la calidad del servicio educativo a través del ejercicio de la suprema inspección y vigilancia de la Educación Superior. (Artículo 3°) En esta ley se especifica que toda persona puede acceder a la educación superior si demuestra que posee las capacidades requeridas y cumple con los requisitos y condiciones académicas exigidas en las Instituciones. Adentrando en los artículos

28, 29 y 30 los cuales reconocen la autonomía de las Universidades, se enfatiza en los derechos que las Instituciones poseen a la hora de crear, organizar y desarrollar programas académicos, así como su respectivo funcionamiento. Así mismo las universidades deben establecer regímenes y aplicar los recursos para el cumplimiento de su misión social, es así como una parte de estos recursos deben ser destinados a la financiación de programas que garanticen la permanencia y graduación de los estudiantes universitarios.

Por otra parte, actualmente se solicita a las Instituciones de Educación Superior su acreditación Institucional y la acreditación de los diferentes programas. Según el artículo 53° de la ley 30 de 1992 se crea el Sistema Nacional de Acreditación para las instituciones de Educación Superior cuyo objetivo fundamental es garantizar a la sociedad que las instituciones que hacen parte del sistema cumplen los más altos requisitos de calidad y que realizan sus propósitos y objetivos. Las Universidades deben acogerse a este proceso de manera voluntaria, en el cual debe realizar una autoevaluación que conlleve a la elaboración de planes de mejoramiento con miras a la excelencia educativa.

Según la ley 30 en los Artículos 117-119 plantea que las Universidades deben generar programas y estrategias de bienestar, entendidos como actividades encaminadas a orientar el desarrollo físico, psicoafectivo y social de los estudiantes, las cuales generan mejores condiciones que contribuyen a la retención de la población estudiantil. Cada institución de Educación Superior destinará por lo menos el dos por ciento (2%) de su presupuesto de funcionamiento para atender adecuadamente su propio bienestar universitario.

El Consejo Nacional de Acreditación ha creado los lineamientos para la acreditación de programas pregrado (CNA, 2013) en el cual se plantea dentro del factor a evaluar denominado “Bienestar Institucional” en la característica N° 32 La Permanencia y retención estudiantil, que establece:

El programa ha definido sistemas de evaluación y seguimiento a la permanencia y retención y tiene mecanismos para su control sin detrimento de la calidad. El tiempo promedio de permanencia de los estudiantes en el programa es concordante con la calidad que se propone alcanzar y con la eficacia institucional.



Aspectos a evaluar:

- Tasas de deserción estudiantil acumulada y por períodos académicos, acorde con los reportes efectuados al Sistema para la Prevención de la Deserción de la Educación Superior – SPADIES–.
- Registros periódicos de la caracterización de los estudiantes teniendo en cuenta variables de vulnerabilidad.
- Existencia de proyectos que establezcan estrategias pedagógicas y actividades extracurriculares orientadas a optimizar las tasas de retención y de graduación de estudiantes en los tiempos previstos, manteniendo la calidad académica del programa.

Se puede evidenciar que uno de los parámetros e indicadores de calidad establecidos como criterio para la evaluación es precisamente la retención de estudiantes y las políticas para enfrentar la deserción, por ende es de gran importancia realizar una intervención en este sentido.

## 2.4 Antecedentes

### *A nivel Internacional*

Los altos índices de estudiantes universitarios que inician una carrera y por diferentes motivos no logran culminarla, convierten la deserción en un grave problema del sistema educativo que no solo afecta a nuestro país sino también a muchos países del mundo. Debido a esto, se han realizado varias investigaciones basadas en enfoques teóricos que han generado resultados importantes los cuales contribuyen a la determinación de los múltiples factores asociados a la deserción.

El estudio titulado: Perfil del alumno desertor en la Universidad Católica de la Santísima Concepción. Un estudio de caso, desarrollado por Magdalena Saldaña en Chile (2009), tiene como objetivo no solo identificar qué factores afectan la permanencia de los estudiantes en la carrera escogida, sino también modelar los datos a través de una regresión logística binaria. Al

igual que en el presente estudio Saldaña trabajó con una serie de variables independientes categóricas y numéricas que influyen sobre una variable dependiente dicotómica (desertor/regular), construyó la base de datos compuesta por 329 alumnos, para obtener una primera aproximación a la información contenida en ella realizó pruebas estadísticas bivariadas entre ellas el chi-cuadrado para observar relaciones entre variables, trabajando con un nivel de significación de  $p \leq .05$ .

Para construir el modelo que identifique las variables que más impactan sobre los eventos de deserción, analizó los datos mediante la Regresión Logística pues considera que es uno de los procedimientos estadísticos más apropiado y además el más utilizado en los modelos de deserción. De esta manera, se generó un modelo multivariado predictivo el cual cumple con las exigencias de parsimonia al contar con variables que predicen correctamente el 90,3% de los casos. Variables como: Ingresos familiares, lugar de residencia, escolaridad previa, promedio académico de los estudiantes, cantidad de créditos aprobados en cada semestre, porcentaje financiero que obtuvo para cubrir costos de matrícula, tuvieron un efecto significativo que permite predecir el evento de deserción. De estas variables el ingreso bruto familiar y el promedio semestral del alumno son las que muestran efectos más fuertes, indicando una relación inversamente proporcional con la probabilidad de deserción (a menores ingresos o rendimiento más bajo, mayor probabilidad de desertar).

El proceso metodológico utilizado por Saldaña es un gran guía de orientación para el presente estudio en la construcción de un modelo estadístico que clasifica a los estudiantes de acuerdo a la probabilidad de deserción, la creación de la base de datos también incluyó a los alumnos de un programa en específico, así mismo ordena los factores que impactan en el fenómeno en 4 grandes grupos y muchas de las variables utilizadas en este trabajo también serán analizadas en la presente investigación.

### ***A nivel nacional***

En Colombia son pocas las investigaciones que se han llevado a cabo para estudiar y cuantificar los distintos factores por los cuales un estudiante deserta, uno de estos estudios fue desarrollado por Vásquez, Castaño, Gallón, & Gómez (2003), el cual se titula: “Determinantes de la deserción estudiantil en la Universidad de Antioquia”, el cual tuvo como objetivo analizar

la deserción en los estudiantes universitarios, particularmente en la facultad de Ingeniería aplicando métodos estadísticos y econométricos que permiten establecer los principales determinantes de este fenómeno con el fin de diseñar políticas que lo controlen.

Para la investigación sus autores tomaron los datos del Sistema de Información de Matrícula y Registro, el análisis se llevó a cabo sobre una muestra de 138 estudiantes. Aunque el estudio se desarrolló con la metodología de análisis de modelos de duración, sirve como base para la presente investigación en la identificación de las variables para los cuatro conjuntos de factores que determinan principalmente la deserción. Una de las recomendaciones metodológicas de los autores es que se deben incluir los cuatro conjuntos de factores como determinantes, ya que considerar factores aislados o la combinación de algunos de ellos no permite obtener conclusiones robustas sobre el evento.

Como resultado se obtuvo que todas las variables incluidas como factores personales e institucionales parecen tener un impacto importante en el riesgo de deserción de los estudiantes. De las variables académicas se concluye que el riesgo de deserción para los buenos estudiantes, es decir aquellos que obtienen un buen promedio crédito por semestre parece ser menor. Sin embargo, el riesgo de deserción aumenta para estudiantes que provengan de colegios privados y que no hayan recibido orientación profesional previo ingreso a la Universidad. Finalmente, en cuanto a las variables socioeconómicas es posible concluir que el nivel educativo de los padres influye en el riesgo de deserción, así no tener padres con nivel educativo bajo (es decir, tener padres profesionales) aumenta muy poco el riesgo de desertar. De mismo modo la responsabilidad de personas a cargo y de responder económicamente por sí mismo aumentan el riesgo de deserción del estudiante.

Otro estudio relacionado con la deserción estudiantil es el desarrollado por Gómez, Padilla, & Rincón (2016) denominado: “Deserción escolar de adolescentes a partir de un estudio de corte transversal: Encuesta Nacional de Salud Mental Colombia 2015”, el cual tiene por objetivo evaluar la relación entre las características individuales y del hogar y la deserción escolar de adolescentes de 12 a 17 años en Colombia.

Esta investigación se basa en la Encuesta Nacional de Deserción donde reporta las causas en las cuales se describen motivos económicos, familiares, individuales y de las instituciones

educativas. La investigación es observacional de corte transversal, estratificado por sexo y por las diferentes regiones del país, a partir de una muestra de 15.351 personas. Algunas variables observadas fueron: edad, sexo, número de personas por hogar, nivel de educación del jefe del hogar, si ha tenido hijos, zona (rural/urbana), nivel de pobreza, entre otras. Este estudio es un orientador para la presente investigación en cuanto a la metodología utilizada, inicialmente calcula frecuencias absolutas y relativas de cada una de las variables incluidas en el estudio, luego se estima la razón de disparidades (odds ratio-OR) para deserción escolar por cada categoría de variables. Se evaluó la asociación entre cada variable y la escolarización a partir de la prueba de  $\chi^2$  de Pearson. Con las variables que en la prueba de asociación obtuvieron  $p < 0,2$ , se realiza un modelo de regresión logística reportando como medida del efecto su OR (IC95%).

Como resultado se observó un mayor porcentaje de deserción entre las mujeres y las personas que no han tenido hijos. De las variables relacionadas con el hogar se encontraron los porcentajes más altos cuando el nivel educativo del jefe del hogar era bajo, cuando el número de personas en el hogar era igual a 2, entre otras. A partir del modelo de regresión logística se evidencia OR significativas en los grupos de mayor edad, con hijos, desplazados por la violencia en algún momento de la vida, cuando habitan 2 personas por hogar respecto a si viven 5 o más personas por hogar, ser mujer, cuando el hogar se ubica en zona rural o está clasificado en estado de pobreza. Se observó un riesgo 2,97 veces mayor en los estudiantes con un jefe del hogar sin ningún grado de educación frente a los que tienen jefe del hogar universitario o posgraduado.

Un aspecto importante dentro de la presente investigación es la utilización del método estadístico de regresión logística para determinar la probabilidad de deserción en los estudiantes del programa, diversos estudios a nivel nacional han utilizado este procedimiento en diferentes ámbitos, uno de ellos realizado en el año (2013) en la Universidad Nacional por Nelcy Escobar, titulado “Análisis de regresión logística para investigación de mercados”, se plantea como una guía metodológica para la aplicación de la técnica cuantitativa de análisis de regresión logística y se constituye en un marco metodológico que permite entender este recurso estadístico y poder aplicarlo a casos académicos y empresariales por medio del software SPSS.

Para la autora es importante tener en cuenta que cuando la variable dependiente es dicotómica se puede recurrir a la regresión logística, pues ésta estima directamente la probabilidad de ocurrencia de un acontecimiento en el que la variable dependiente toma dos

categorías. Esta técnica puede usarse para fines explicativos, donde se pretende analizar la relación de una variable con otras; o para fines predictivos, que es cuando permitiría estimar el comportamiento de una variable en función de lo que se sepa de otras variables que influyen su comportamiento. De aquí, que esta técnica estadística contribuya en las ciencias sociales para determinar cuáles variables tienen que ver con el desarrollo de un suceso, cuáles variables afectarán la toma de una decisión, o cuáles variables permitirán predecir la ocurrencia de un evento.

Esta investigación sirve como guía técnica para realizar el análisis estadístico, en primer lugar se realiza la estadística descriptiva para cada una de las variables, luego se hacen tablas de contingencia para estimar la relación entre las variables independientes y la variable dependiente, posteriormente se realiza el análisis de regresión logística, al obtener la ecuación realiza pruebas reemplazando los valores de las variables independientes, además analiza los Odds ratio los cuales indican cómo incrementa la probabilidad de ocurrencia frente a la de no ocurrencia. Según los resultados de los 447 casos en total el modelo logró clasificar correctamente el 79.7%.

### *A nivel regional*

En el ámbito regional, en la Universidad de Nariño y en la Institución Universitaria CESMAG, se desarrolló un proyecto de investigación interinstitucional realizado por Timarán & Jiménez (2015), denominado: “Extracción de perfiles de deserción estudiantil en la institución universitaria CESMAG”, cuyo objetivo fue detectar patrones de deserción estudiantil a partir de los datos socioeconómicos, académicos, disciplinarios e institucionales de los estudiantes de los programas de pregrado de estas dos instituciones. Esta investigación acogió como metodología las diferentes etapas del proceso de descubrimiento de conocimiento en bases de datos, se tomó la información correspondiente a los años 2004, 2005 y 2006 con el fin de hacerles un seguimiento completo hasta el año 2011, para determinar si desertaron o no, finalmente se obtuvieron perfiles de deserción estudiantil con la ayuda de la técnica de clasificación basada en árboles de decisión.

Este estudio realiza un importante aporte a la presente investigación en cuanto a la selección de variables que se encuentran disponibles en las bases de datos de la oficina de OCARA y la construcción del perfil del estudiante desertor, Como resultado se obtuvo la

clasificación con una confianza mayor que 75%, si el promedio de notas es menor que 2,4, el estudiante deserta, lo mismo ocurre si el promedio de notas está entre 2,4 y 3,1. Por consiguiente, los factores predominantes en la deserción estudiantil son los académicos, especialmente un promedio de calificaciones bajo. Dentro de los factores socioeconómicos también predomina el costo de la matrícula, tener 18 años al ingresar a la institución, pertenecer al estrato 1 ó 3, y ser originario de la Costa Pacífica o de la ciudad de Pasto.

Otro estudio se realizó en la facultad de Ciencias Exactas y Naturales de la Universidad de Nariño, al cual pertenece el programa de Licenciatura en informática, denominado: Análisis de la deserción estudiantil en la facultad de ciencias exactas y naturales de la Universidad de Nariño desde la cohorte 2001-2 hasta la cohorte 2006-2 utilizando el sistema SPADIES, de los autores Calpa & Cruz (2008), permite identificar las diferentes características de los estudiantes desertores.

La aplicación SPADIES se basa en un modelo estadístico de duración, esta herramienta hace seguimiento a las cifras de deserción en estudiantes de educación superior con datos que suministran las instituciones para agrupar a los estudiantes de acuerdo con su riesgo de deserción. Una vez que se ingresan los datos al sistema, se realiza una sincronización con el servidor del Ministerio de Educación Nacional para contar con esta información a nivel nacional y permitir realizar consultas que se devuelven en tablas de cantidades, gráficas y porcentajes donde pueden ser analizadas y comparados por el usuario. Es así como permite conocer la evolución y rendimiento académico de los estudiantes, lo cual permite establecer los factores que inciden en la deserción y estimar el riesgo de deserción de cada estudiante.

Esta investigación incluyó en total 1440 estudiantes, donde se analizó e interpretó los motivos reales que impulsan la deserción estudiantil, El programa de licenciatura en informática se destaca por una tasa de deserción inferior a los demás programas de la facultad según el estudio, en cuanto al género la tasa de deserción en los hombres es mayor. Las personas que trabajan antes de presentar las pruebas de estado tienden a desertar con mayor frecuencia, con un porcentaje de 55,83%. Con respecto a la variable edad de presentación del ICFES es muy marcada la deserción con un 75% en estudiantes de 26 años o más al momento de presentar la prueba, otra variable muy importante para el presente proyecto es el ingreso familiar, donde la mayor tasa de deserción se encuentra en los estudiantes cuyos ingresos familiares oscilan entre 0

y 1 salario mínimo, con 58,04%, el índice de deserción se reduce a medida que los ingresos aumentan.

## 2.5 Marco de Contextual

La Universidad de Nariño fundada en 1904, es una entidad de educación superior con acreditación institucional de alta calidad ubicada en el municipio de Pasto, cuenta con 11 facultades, 91 programas con registro calificado, 22 de ellos están acreditados. Esta institución desde su creación tiene como principal objetivo ofrecer una educación completa a todos los ciudadanos de la región sur de Colombia, con el compromiso de formar profesionales que sean totalmente competentes y comprometidos con el desarrollo de la región, así como del país.

La Facultad de Ciencias Exactas y Naturales de la Universidad de Nariño desde 2007, es la unidad Académico-Administrativo constituida por el Departamento de Matemáticas y en él, mediante Acuerdo 145 de octubre 3 de 1992 del Honorable Consejo Superior de la Universidad de Nariño fue creado el Programa de Licenciatura en Informática adscrito a esta Facultad.

Según la información encontrada en página web de la Universidad de Nariño la misión y visión del programa de Licenciatura en informática son:

**Misión:** El Programa de Licenciatura en Informática, acorde con los lineamientos de la Universidad de Nariño, tiene como misión preparar talento humano en los niveles de educación superior y no formal, con formación integral y excelencia académica en las áreas de Informática y Pedagogía; capaces de desempeñarse con idoneidad y competitividad, y que mediante la adecuada utilización de los medios informáticos en los espacios técnico, científico, administrativo y cultural, propendan por la búsqueda de soluciones adecuadas para los problemas educativos de su entorno a través de la investigación, la docencia y la proyección social.

**Visión:** El Programa de Licenciatura en Informática de la Universidad de Nariño, se proyecta como una propuesta curricular de alta calidad para la formación integral de licenciados capaces de contribuir al desarrollo socio-económico, político y cultural a través de la gestión de soluciones en las áreas de informática y pedagogía, con miras a alcanzar el

liderazgo regional y nacional en la utilización de la informática y las nuevas tecnologías de la información y la comunicación.

**Composición de la población:** “El programa de licenciatura en Informática de la Universidad de Nariño ha contado con un promedio de 248 estudiantes Inscritos en el periodo académico 2010 – 2015” (Udenar en cifras, pág. 32). El programa recibe a estudiantes de diferentes sectores del sur del país, con diversidad de características para formar docentes capaces de brindar a los estudiantes las herramientas necesarias para el aporte hacia el desarrollo y la tecnología.

Según el plan de mejoramiento del programa rescatado del proceso de acreditación institucional, faltan registros periódicos de estudiantes en situación de vulnerabilidad y estrategias de vinculación de estudiantes en redes de apoyo frente este tipo de situaciones, debido a su caracterización y mayor riesgo de deserción es innegable la necesidad de contar con esta y todo tipo de información que brinde los caminos necesarios para plantear estrategias adecuadas.

### **3. Procedimiento Metodológico**

#### **3.1 Aspectos investigativos**

##### **3.1.1 Enfoque.**

Esta investigación tiene un enfoque cuantitativo, representado en el procesamiento de datos estadísticos que permite establecer patrones de comportamiento asociados al fenómeno de estudio. Para Galeano (2004):

Los estudios de corte cuantitativo pretenden la explicación de una realidad social vista desde una perspectiva externa y objetiva, su intención es buscar la exactitud de mediciones o indicadores sociales con el fin de generalizar sus resultados a poblaciones o situaciones amplias, trabajan fundamentalmente con el número, el dato cuantificable. (pág. 23)



Los datos se analizan por medio de procedimientos estadísticos, es decir la medición se transforma en datos cuantificables que serán considerados a través de la estadística logrando inferencias basadas en el cálculo de probabilidades.

### **3.1.2 Alcance.**

Esta investigación tiene un alcance correlacional, este tipo de estudios pretenden conocer la asociación entre las variables que intervienen en el fenómeno de estudio, con el objetivo de evaluar el grado de relación entre ellas. El alcance correlacional se caracteriza porque primero se miden las variables y luego con la aplicación de técnicas estadísticas se estima la correlación. Esto permite consolidar las bases para predecir con mayor exactitud la probabilidad de deserción según el valor de las variables.

La utilidad principal de los estudios correlacionales es saber cómo se puede comportar un concepto o una variable al conocer el comportamiento de otras variables vinculadas, es decir, intentar predecir el valor aproximado que tendrá un grupo de individuos o casos en una variable, a partir del valor que poseen en la o las variables relacionadas. (Hernández, Fernández, & Baptista, 2006, pág. 82)

### **3.1.3 Diseño.**

Se utiliza un diseño no experimental que tiene como propósito analizar los factores que inciden en la deserción, con el fin de determinar la probabilidad de riesgo de un estudiante a partir de la información obtenida de OCARA y de base de datos externas, es decir no hay manipulación en los datos, estos son analizados tal y como han sido entregados.

#### ***3.1.3.1 Operacionalización de Variables.***

Para la interpretación de la información se establece los determinantes de la deserción estudiantil para efectos de estudio se ha tomado la siguiente distribución de las variables

Tabla 5

*Operacionalización de Variables*

Factor	Variable	Definición Conceptual	Indicador
<b>Factor Individuales</b>	<b>Edad Ingreso</b>	Edad del estudiante al ingresar a la institución.	Registro OCARA
	<b>Sexo</b>	Grupo al que pertenecen los seres humanos de cada género, entendido este desde un punto de vista sociocultural en lugar de exclusivamente biológico. (RAE, 2014)	Registro OCARA
	<b>Estado Civil</b>	Condición de una persona en relación con su nacimiento, nacionalidad o matrimonio, que se hacen constar en el registro civil. (RAE, 2014)	Registro OCARA
	<b>Régimen de Salud</b>	Régimen de salud al que está afiliado el estudiante.	Registro OCARA
	<b>Zona Residencia</b>	Zona del Departamento de Nariño o del país donde reside el estudiante al ingresar en la institución.	Registro OCARA
<b>Factor Académico</b>	<b>Tipo de Colegio secundaria</b>	Tipo de institución donde realizó la secundaria.	Registro OCARA
	<b>Puntaje del Icfes</b>	Valor promedio de las áreas de la prueba del ICFES.	Registro OCARA
	<b>Promedio de Notas</b>	Promedio de notas del estudiante.	Registro OCARA
	<b>Materias Perdidas</b>	Número de materias que ha perdido el estudiante en la carrera.	Registro OCARA
	<b>Área de Materia</b>	Área a la que pertenecen la mayoría de materias perdidas.	Registro OCARA
	<b>Veces Perdida</b>	Número de veces que ha perdido una materia.	Registro OCARA

	<b>Deserción</b>	Determina si el estudiante desertó o no.	Registro OCARA
<b>Factores Socioeconómicos</b>	<b>Estrato Social</b>	Un estrato social está constituido por un conjunto de personas, agregados sociales, que comparten un sitio o lugar similar dentro de la jerarquización o escala social, donde comparten similares creencias, valores, actitudes, estilos y actos de vida. (Casadesús, 2013, pág. 56)	Registro OCARA
	<b>Tipo de Residencia</b>	Variable en la que se establece si el estudiante vive en una residencia propia o arrendada.	Registro OCARA
	<b>Número de hermanos</b>	Aquí se afirma la cantidad de hermanos que tiene un estudiante.	Registro OCARA
	<b>Personas a cargo</b>	Número de personas a cargo	Registro OCARA
	<b>Ocupación de la Madre</b>	Hace referencia a la actividad económica que realiza la madre.	Registro OCARA
	<b>Nivel educativo de la Madre</b>	Se refiere al grado más alto de estudios que la madre del estudiante ha cursado y ha abandonado ya, bien porque los haya concluido o bien porque los interrumpió antes de terminarlos.	Registro OCARA
	<b>Ocupación del Padre</b>	Hace referencia a la actividad económica que realiza el padre.	Registro OCARA
	<b>Nivel educativo del Padre</b>	Se refiere al grado más alto de estudios que el padre del estudiante ha cursado y ha abandonado ya, bien porque los haya concluido o bien porque los interrumpió antes de terminarlos.	Registro OCARA
	<b>Situación laboral del estudiante</b>	El estudiante al momento se encuentra cumpliendo con un trabajo determinado.	Registro OCARA
<b>Valor de la Matrícula</b>	Valor promedio de la matrícula pagada por el estudiante durante la carrera	Registro OCARA	

	<b>Ingresos de la Familia</b>	Determina el valor en salarios mínimos de los ingresos en el núcleo familiar.	Registro OCARA
	<b>Vive con la Familia</b>	Si el estudiante vive con la familia o no	Registro OCARA
<b>Factores Institucionales</b>	<b>Becas y formas de financiamiento</b>	El estudiante obtiene ayudas de financiamiento. (Monitorias, Tutoría, Becas de alimentación, Subsidio de Vivienda)	Registro OCARA
	<b>Sanciones Disciplinarias</b>	El estudiante tiene algún tipo de sanción impuesta por la institución	Registro OCARA

Fuente: Esta investigación

### 3.1.4 Población y Muestra

En esta investigación la población está constituida por los estudiantes desertores y no desertores de los diferentes semestres del programa de Licenciatura en Informática.

La muestra está compuesta por la cohorte establecida entre 2012-1 a 2016-2. Para Lazcano, Fernández, Salazar, & Hernández (2000):

Los estudios de cohorte se han utilizado de manera clásica para determinar la ocurrencia de un evento específico en un grupo de individuos inicialmente libres del evento en estudio. Es decir, toda la población en estudio se sigue a través del tiempo y se compara la incidencia del evento de estudio en individuos expuestos con la de los no-expuestos. (pág. 230)

### 3.1.5 Métodos de recolección de la información.

Para esta investigación es necesario solicitar la información académica, socioeconómica y personal de los estudiantes del programa de Licenciatura en Informática, en la oficina de Control de Admisiones y Registro Académico OCARA. Para complementar los datos faltantes se consultan fuentes externas como las bases de datos del Instituto Colombiano para el Fomento de la Educación Superior (ICFES), del Departamento Administrativo Nacional de Estadística (DANE), del Sistema de Identificación de Beneficiarios Potenciales de Programas Sociales (SISBEN) e información de la Registraduría Nacional del Estado Civil Colombiano.

### **3.2 Diseño Procedimental**

La presente investigación toma como base los datos individuales, socioeconómicos, académicos e institucionales de los estudiantes del programa de Licenciatura en informática, esta información será sometida a un seguimiento completo que permita definir si el estudiante finalmente desertó o no del programa y si lo hizo, identificar cuáles fueron los factores que originaron este hecho, posibilitando la elaboración de un perfil del estudiante desertor.

La construcción del perfil conlleva a la detección e interpretación de los factores más relevantes que se asocian en la deserción estudiantil, a esta información se le aplica el modelo de regresión logística que permite calcular el riesgo de deserción de los estudiantes actuales por medio de la sistematización de un modelo estadístico de clasificación.

#### **3.2.1 Etapa de Selección de Datos.**

El objetivo de esta etapa es obtener las fuentes internas de datos que sirven de base para la construcción del perfil del desertor. Como fuente interna en la Universidad de Nariño se selecciona la base de datos de la Oficina de Control de Admisiones y Registro Académico (OCARA). Teniendo en cuenta la cohorte establecido para la investigación (2012-1 a 2016-2), en esta base de datos se encuentra almacenada la información académica y personal de los estudiantes pertenecientes al programa.

Para complementar los datos faltantes se consultan como fuentes externas las bases de datos del Instituto Colombiano para el Fomento de la Educación Superior (ICFES), del Departamento Administrativo Nacional de Estadística (DANE), del Sistema de Identificación de Beneficiarios Potenciales de Programas Sociales (SISBEN) e información de la Registraduría Nacional del Estado Civil Colombiano.

#### **3.2.2 Etapa de Procesamiento de Datos.**

El objetivo de esta etapa es obtener los datos limpios, sin valores nulos que permitan adquirir patrones de calidad, con la información obtenida se hacen los procedimientos de reducción, recodificación, integración y de limpieza respectiva de los mismos, para evitar redundancias y datos repetitivos. Se identifica la variable dependiente por cada uno de los estudiantes de acuerdo con las variables independientes tal cual se presentan.

Posteriormente se sintetiza la información por medio de un programa informático IBM-SPSS versión 24 que es una herramienta de tratamiento de datos que brinda un resumen estadístico y tablas de frecuencia, con la cual se establece las tendencias en cuanto a las variables, obteniendo la información más relevante relacionada a los diferentes factores que llevaron a los estudiantes a tomar la decisión de desertar.

### **3.2.3 Aplicación Modelo de regresión logística.**

El análisis de regresión logística persigue cuantificar la importancia de la relación existente entre cada una de las variables independientes (también llamadas covariables) y la variable dependiente, y clasificar a los individuos dentro de las categorías de la variable dependiente según la probabilidad que tenga de pertenecer a una de ellas dada la influencia de las covariables. (López & Fachelli, 2016, pág. 6)

El resumen estadístico permite lograr las primeras inferencias asociadas al fenómeno en cuestión, obteniendo estimaciones de características numéricas y asociaciones entre variables. El modelo es útil en este tipo de investigaciones en que la respuesta puede tomar únicamente dos valores 1 presencia y 0 ausencia, además el perfil de variables puede estar formado por caracteres cuantitativos y cualitativos, y se pretende hacer intervenir a todos en una única ecuación.

Para López & Fachelli (2016) el proceso de análisis de una regresión logística se puede dividir en varias etapas:

1. *Selección de las variables del modelo:* hay que determinar, por un lado, el nivel de asociación entre cada variable independiente o explicativa por separado y la variable dependiente, lo que nos permitirá descartar aquellas variables que empíricamente no manifiestan una relación con la variable a explicar y nos puede sugerir asimismo posibles agrupaciones de valores de las variables.
2. *Estimación de los coeficientes de las variables independientes:* la estimación de los coeficientes o pesos de la ecuación de regresión que determinan la importancia de cada variable independiente en la explicación de la dependiente se realiza mediante un algoritmo iterativo de máxima verosimilitud propio del modelo de la regresión logística.

3. *Clasificación de los casos:* en función de la ecuación de regresión logística estimada se procede a la clasificación de los individuos según la variable dependiente pronosticada (pág. 7).

A partir de los coeficientes de regresión ( $\beta$ ) de las variables independientes introducidas en el modelo se establece la fuerza de asociación entre ellas y se puede obtener directamente la OR (odds Ratio) la cual corresponde a la probabilidad de deserción en un individuo con un valor  $x$  frente a la probabilidad que tiene un individuo con un valor  $(x-1)$ . Por ejemplo si  $X$  es la variable EDAD (en años cumplidos) la OR será la probabilidad de deserción que tiene una persona de 25 años en relación a otra que tiene 24 años. Para Aguayo (2007):

Lo que se pretende mediante el modelo de regresión logística es expresar la probabilidad de que ocurra el evento en cuestión como función de ciertas variables, que se presumen relevantes o influyentes, por tanto la ecuación general es:

$$P(Y=1) = \frac{1}{1 + \exp(-\alpha - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \dots - \beta_K X_K)}$$

Posteriormente se realiza una prueba de bondad de ajustes denominada ómnibus la cual prueba más de dos variables independientes, si la significación es menor de 0,05 indica que el modelo ayuda a explicar el evento, es decir, las variables independientes explican la variable dependiente, permitiendo el ajuste del modelo.

### **3.2.4 Etapa de Sistematización del Modelo Estadístico.**

Para optimizar y concretar los resultados obtenidos se construye una aplicación web en la cual se calcula la probabilidad de deserción de un estudiante, de acuerdo con los valores de las variables predictoras en un caso específico.

#### **3.2.4.1 Metodología para el desarrollo de la aplicación web.**

Para el diseño de una aplicación web que calcule la probabilidad de deserción, se utiliza la metodología IWeb (Ingeniería Web) definida como: “el proceso con el que se crean WebApps de alta calidad” (Pressman, 2005, pág. 503), porque permite satisfacer la necesidad de una aplicación con técnicas sistemáticas, estándares y garantías de calidad. Es una metodología

enfocada a la producción, implantación y mantenimiento de aplicaciones y sistemas web caracterizados principalmente porque en su elaboración se utilizan diversidad de herramientas de ingeniería.

### *Etapas del Proceso de la Metodología IWeb*

La metodología IWeb está constituida por las siguientes etapas:

1. **Formulación:** En esta etapa se identifican los objetivos y metas para la producción de la aplicación web, donde es necesario determinar y observar las necesidades y tareas que pueden mejorar en su organización y establecer quién va a utilizarla.
2. **Planificación:** En la etapa de planificación se estima el costo global del proyecto y riesgos del mismo, como también la planificación empleada a lo largo del proceso IWeb.
3. **Análisis:** Consiste en establecer los requerimientos técnicos y requisitos de diseño además se determinan los elementos de contenido, funcionalidad e interacción que se incorporan en la aplicación web.
4. **Ingeniería:** En esta etapa de la metodología IWeb se recopila la información y se diseña el contenido, seguido del diseño de la arquitectura, navegación e interfaz del usuario.
5. **Generación de Páginas:** Es la etapa de construcción donde se unen los diseños arquitectónico, de navegación y de interfaz para elaborar las páginas Web en HTML5 y CSS.
6. **Pruebas y Evaluación del cliente:** Para esta etapa se revisa la aplicación con el fin de encontrar y corregir errores. Se realizan diferentes pruebas de funcionamiento en cada una de las etapas anteriores con el fin de presentar la plataforma funcional en todos los aspectos al usuario final.

Finalmente, se propone como estrategia la utilización del sistema para determinar el riesgo de deserción en los estudiantes del programa.



## 4. Resultados

### 4.1 Selección de Datos

De acuerdo con los datos obtenidos de las fuentes de información, se toman los grupos de factores que a continuación se presentan con sus respectivas variables.

**Tabla 6**

*Variables obtenidas en la selección de datos*

<b>Grupo de Factores</b>	<b>Variables</b>
<b>Factores Individuales</b>	Sexo
	Edad de Ingreso
<b>Factores Académicos</b>	Tipo de colegio
	Puntaje Icfes
	Promedio de Notas
	Último semestre cursado
	No. Materias perdidas último semestre
	No. Materias perdidas en total
<b>Factores Socioeconómicos</b>	Estrato Social
	Ciudad de Procedencia
	Lugar de Residencia
	Tipo de Residencia
	Vive con la Familia
	Tiene Hermanos en la U
	Jefe de Familia
	Ingresos Familiares Anuales
	Puntaje socioeconómico U
	Valor de la Matrícula
<b>Factores Institucionales</b>	Tipo de cupo
	Sanciones

---

Fuente: Esta investigación

## 4.2 Procesamiento de Datos

### 4.2.1 Análisis Univariado.

Una vez concluidas las etapas de recolección y procesamiento de datos, se inicia con una de las fases más importantes de esta investigación: el análisis de datos. Se examina descriptivamente las variables de modo independiente, abarcando el periodo comprendido entre el primer semestre del año 2012 hasta el último semestre de 2016, se realiza un análisis univariado para identificar los comportamientos generales de la población.

#### 4.2.1.1 Distribución de Frecuencias.

En el estudio la muestra está conformada por 275 estudiantes, la distribución de frecuencias indica el número de casos que hay en cada categoría de la variable, a partir de estos valores se calcula el porcentaje con respecto al total de la muestra.

#### *Factores Individuales*

**Distribución por género:** Como se puede observar en la figura 3, de acuerdo con los datos obtenidos y tabulados, se encuentra una gran diferencia entre la cantidad de hombres y mujeres. Así el 68,36% de la población es de género masculino y el 31,6% femenino, por lo que se puede considerar que la carrera de Licenciatura en Informática resulta más atractiva para los hombres (Ver anexo 1).

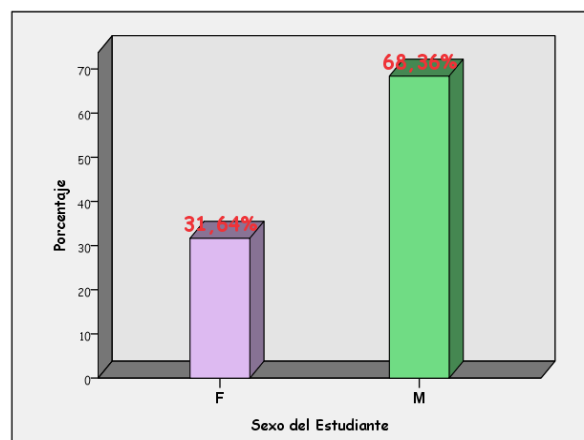


Figura 3. Distribución por sexo del estudiante.  
Fuente: Esta investigación

**Distribución para edad de ingreso:** La variable Edad de ingreso es cuantitativa, por tanto se han creado ciertos rangos que permiten interpretar de mejor manera los resultados. Según la figura 4; el 65.8% de los estudiantes son menores de 20 años, lo que indica que la gran mayoría son muy jóvenes y esto podría llegar a ser un causal de deserción. El 26.5% tienen un rango de edad entre 20 y 24 años, un porcentaje muy pequeño de la población 6.9% tiene entre 25 y 29 años y menos de 1% son mayores de 30 años. Según la Tabla No. 7 la edad promedio es 19.7 años, con una desviación estándar de 4.6 años, un valor mínimo de 16 y un máximo de 32 (Ver anexo 2).

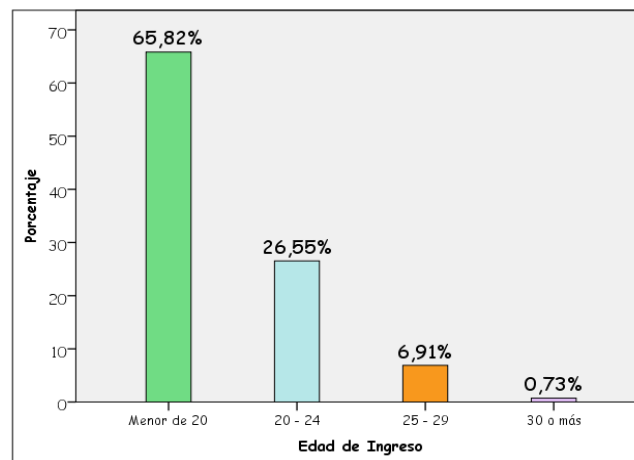


Figura 4. Distribución por edad de ingreso.  
Fuente: Esta investigación

### ***Factores Académicos***

**Distribución por tipo de colegio:** Esta variable indica la clase de colegio en el que el estudiante terminó su bachillerato, como se puede observar en la figura 5, el 89.8% de los estudiantes provienen de un colegio oficial o público y el 9.1% de una institución privada, contrario a lo que se podría pensar sobre los colegios privados que gozan de cierta fama en cuanto a la calidad académica, muchas instituciones públicas han logrado posicionarse y ser reconocidas a nivel regional como las mejores, de ahí que los colegios oficiales puntúan en esta variable (Ver anexo 3)..

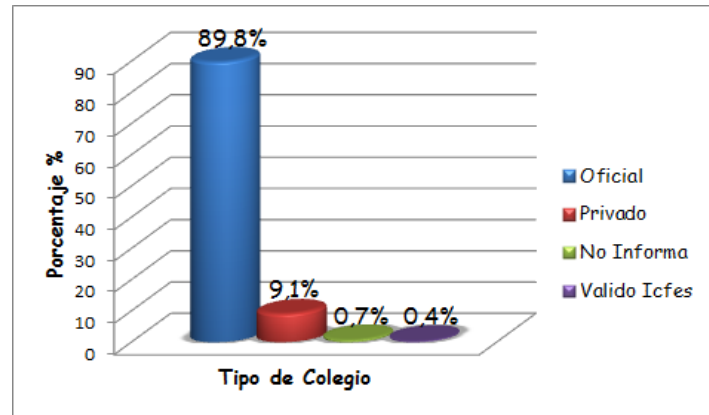


Figura 5. Distribución por tipo de colegio.  
Fuente: Esta investigación

**Distribución por Puntaje Icfes:** Una variable muy importante que puede ser determinante para la deserción es el puntaje icfes, el cual es obligatorio para el ingreso a la educación superior, pues la universidad lo exige como requisito para la admisión. En la figura 6 se observa que el 78.9% de los estudiantes tiene un promedio entre 51 y 60.99 puntos. Según la tabla No. 7, el puntaje promedio es de 56.7, con una desviación estándar de 4.55, un valor mínimo de 42 y un máximo de 77.3 (Ver anexo 4).

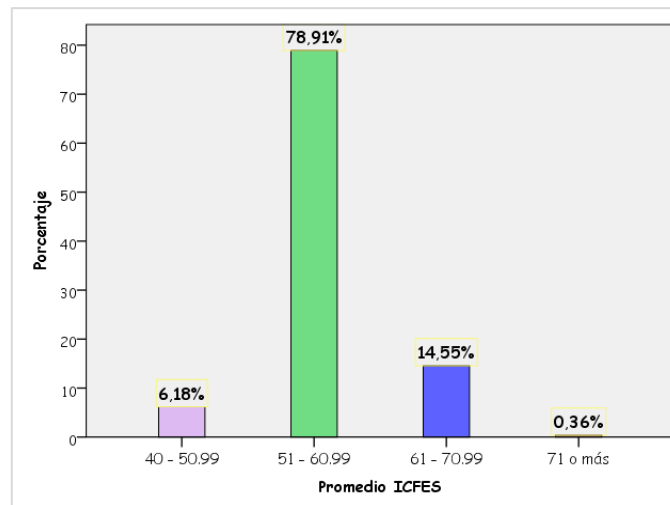


Figura 6. Distribución por puntaje ICFES.  
Fuente: Esta investigación

**Distribución por Promedio de notas:** El promedio de notas es una medida cuantitativa que evalúa el rendimiento académico del estudiante, es un factor que puede incidir en el riesgo de deserción. Según la figura 7, más de la mitad de los estudiantes; 56.3% tiene un promedio

regular entre 3 y 4 puntos, el 15.6% oscila entre 4 y 5 puntos, el resto de la población tiene un promedio menor de 3. La tabla No. 7, indica que el promedio de notas para todos los estudiantes es de 3.2, con una desviación estándar de 0.89, un valor mínimo de 0 y un máximo de 4.65 (Ver anexo 5).

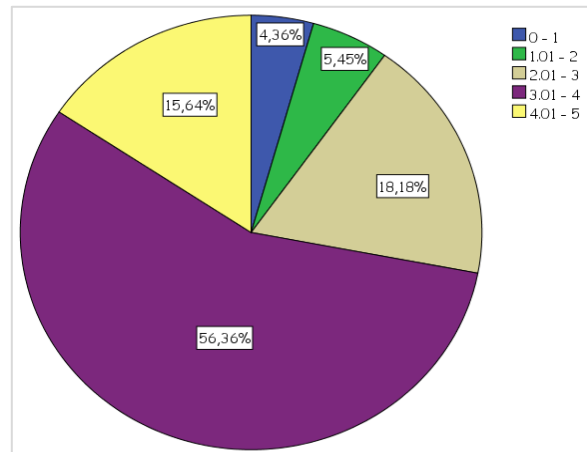


Figura 7. Distribución por promedio de notas.

Fuente: Esta investigación

**Distribución por último semestre cursado:** La figura 8, indica que el primer y tercer semestre son los que más cantidad de estudiantes tienen en el programa con un 20.7% y 23.6% respectivamente, seguidos del quinto y séptimo semestre con un porcentaje similar para los dos del 14%, el noveno y décimo semestre también comparten un porcentaje semejante del 8%. El segundo, cuarto, sexto y octavo semestre son los que menos cantidad de estudiantes tienen al momento de realizar este estudio ya que se encontraban activos los semestres impares (Ver anexo 6).

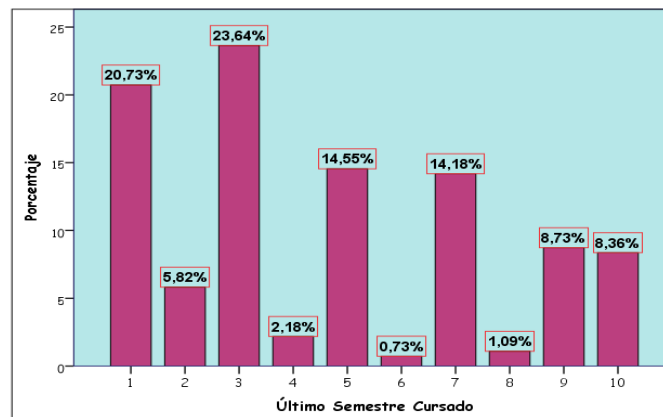


Figura 8. Distribución por último semestre cursado.

Fuente: Esta investigación

**Distribución por No. de materias perdidas en el último semestre:** Al observar la figura 9, se aprecia que el 41.5% de los estudiantes no perdieron ninguna materia en el semestre inmediatamente anterior, mientras que el 18.2% perdieron una asignatura, cerca del 15% perdieron 2 y el resto de estudiantes perdieron más de 3 materias. De ello se puede inferir que más de la mitad de los estudiantes perdieron al menos una asignatura en el último semestre cursado (Ver anexo 7).

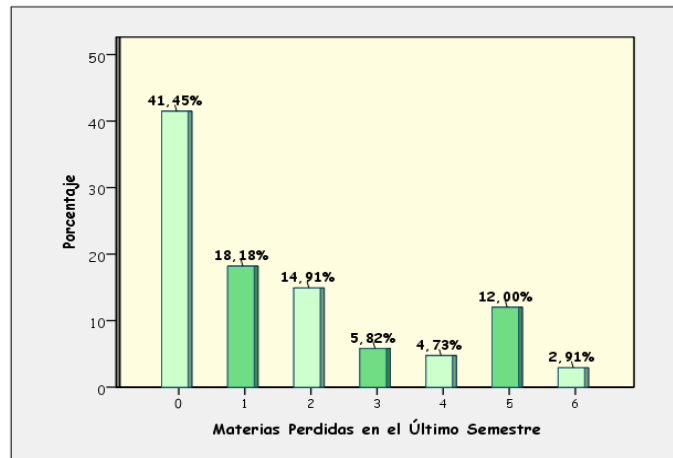


Figura 9. Distribución por materias perdidas en el último semestre.  
Fuente: Esta investigación

**Distribución por Número de materias perdidas en total:** Debido a que la variable es cuantitativa se agrupó en rangos para una mejor interpretación. La figura 10 presenta la información sobre la cantidad de materias que los estudiantes han perdido durante la carrera, un poco más de la mitad han perdido de 0 a 3 materias, el 38.5% entre 4 y 7 asignaturas, el 7.6% de 8 a 11 materias y el 2.5% 12 o más. Según la tabla No. 7, el promedio de materias perdidas en total para todos los estudiantes es de 3.6, con una desviación estándar de 2.9, un valor mínimo de 0 y un máximo de 17 (Ver anexo 8).

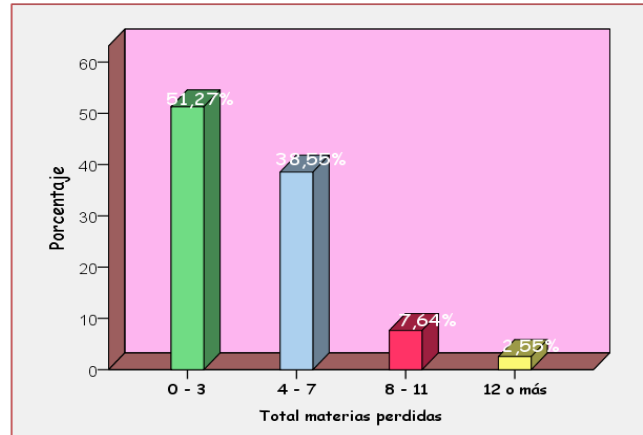


Figura 10. Distribución por total de materias perdidas.

Fuente: Esta investigación

### *Factores Socioeconómicos*

**Distribución por estrato social:** El estrato social indica el nivel socioeconómico de los estudiantes, de la gráfica 11 se puede inferir que la mayoría se encuentra en los estratos más bajos; 1 y 2, con un 48.7% y 43.6% respectivamente, mientras que tan solo el 7.3% son de estrato 3 y el 0.36% de estrato 4. Esto permite argumentar que al ser una Universidad pública alberga estudiantes de bajos recursos que hacen el esfuerzo para ingresar, debido a que el acceso a una universidad privada es muy difícil y costoso. Esta variable puede ser un gran indicador del evento en estudio ya que la situación económica de un estudiante lo puede empujar a abandonar la carrera y desertar (Ver anexo 9).

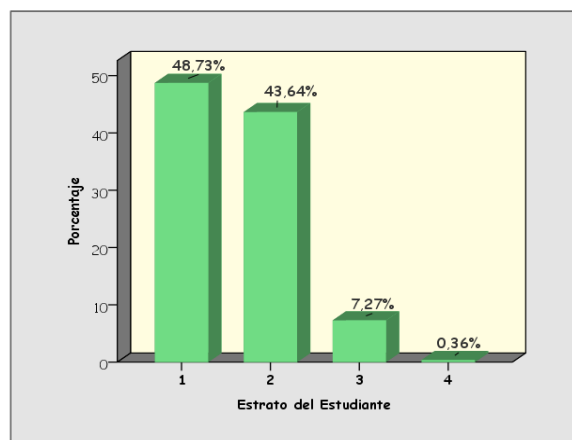
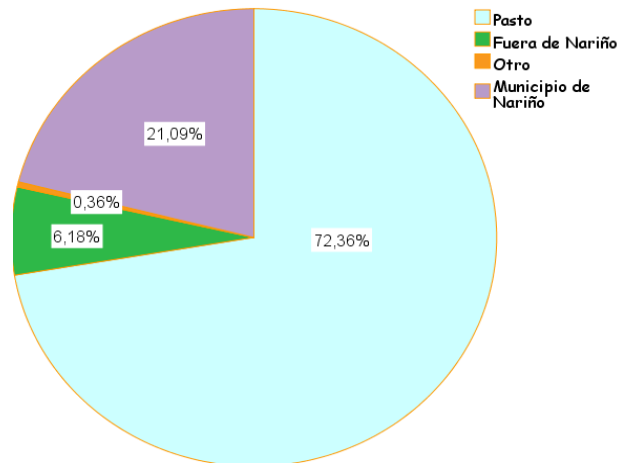


Figura 11. Distribución por estrato.

Fuente: Esta investigación

**Distribución por lugar de procedencia:** Este factor hace referencia a la ciudad o municipio de origen del estudiante, se incluye en este estudio porque se presume es un factor determinante. Según la figura 12 el 72.4% de la población objeto de estudio procede de la ciudad de Pasto, el 21.5% de otros municipios de Nariño, el 6.1% de lugares fuera de Nariño y el 0.3% de otro sitio (Ver anexo 10).



*Figura 12.* Distribución por lugar de procedencia.  
Fuente: Esta investigación

**Distribución por lugar de residencia:** Esta variable indica el lugar donde vive actualmente el estudiante, se divide en 3 categorías para facilitar su estudio. Como puede observarse en la figura 13 gran parte de la población estudiantil del programa 72.7% reside en la ciudad de Pasto, con respecto a la zona rural de Pasto un 12,7% vive en estas localidades y el 14,5% habita en otro municipio (Ver anexo 12).

El lugar de residencia puede considerarse como un factor de riesgo, por ejemplo: los habitantes de zonas apartadas de la ciudad o de otros municipios diariamente se enfrentan a diferentes situaciones que afectan su permanencia en la universidad; como, problemas de movilidad, elevados costos de alimentación y estadía, falta de recursos económicos, tiempo, etc.



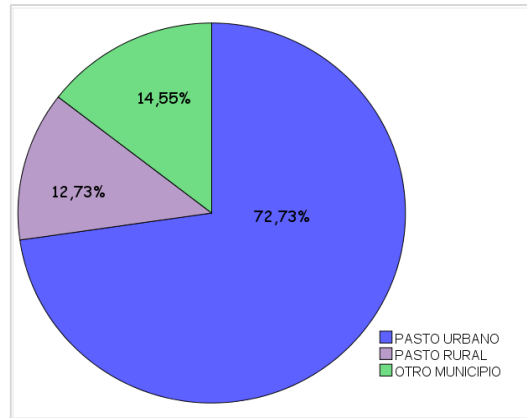


Figura 13. Distribución por lugar de residencia.  
Fuente: Esta investigación

**Distribución por tipo de residencia:** El tipo de residencia hace alusión a la clase de vivienda que habita el estudiante. Según la figura 14, la mayor proporción: 64.7%, tiene vivienda propia, mientras que el 24.7% arrienda o anticresa, el 2.9% de la población tiene vivienda propia pero la paga en cuotas. Esta información es importante porque permite concluir que la mayoría de los estudiantes gozan de un gran beneficio al tener una vivienda propia, lo que mejora su calidad de vida (Ver anexo 12).

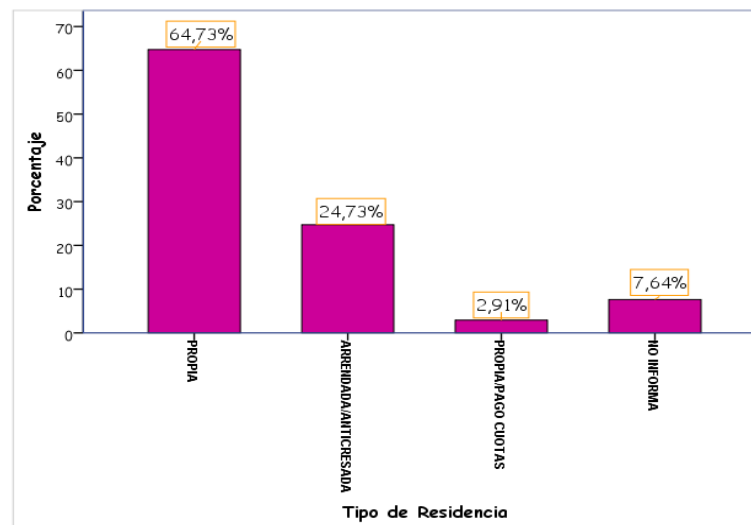
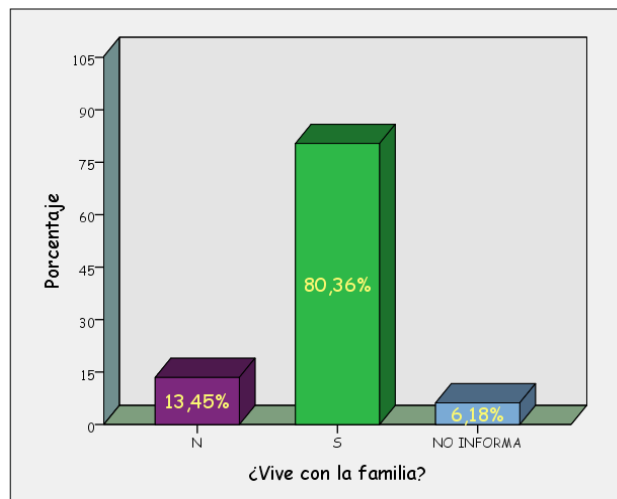


Figura 14. Distribución por tipo de residencia.  
Fuente: Esta investigación

**Distribución por vive con la familia:** Vive con la familia, es una variable que indica si el estudiante reside solo o con su núcleo familiar, es relevante porque señala el grado de dependencia o independencia que tiene la población. Se puede pensar que una persona que vive sola tiene más responsabilidades en cuanto al manejo de su casa y sus recursos, en cambio los que viven con sus familiares reciben apoyo constante, de alguna manera esto podría incidir en el riesgo de deserción. Como se observa en la figura 15, en este caso el 80.4% de los estudiantes viven con la familia, mientras que el 13.5% no, el porcentaje restante no informa (Ver anexo 13).



*Figura 15:* Distribución por vive con la familia.  
Fuente: Esta investigación

**Distribución por hermanos en la universidad:** Según la figura 16, una gran proporción de los estudiantes 85.1% no tiene hermanos en la universidad y el 14.9% si los tiene. Esta variable puede llegar a influir en la deserción, ya que una familia que tenga dos o más hijos en la universidad tendría más dificultades para que ellos terminen su carrera que una en la que solo hay un hijo estudiando (Ver anexo 14).

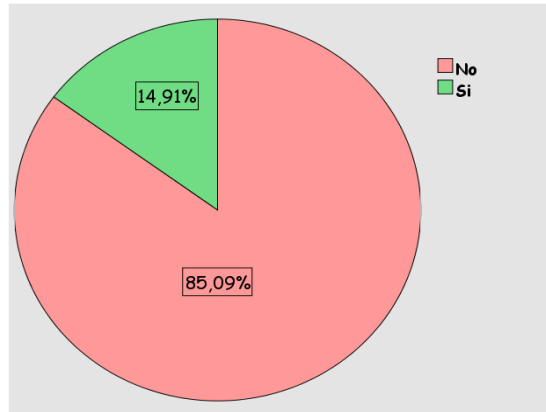


Figura 16. Distribución por hermanos en la universidad.  
Fuente: Esta investigación

**Distribución por jefe de familia:** Esta variable hace referencia a la persona responsable la cual mantiene el hogar del estudiante. En nuestra sociedad es cada vez más frecuente que las madres se vean en la obligación de cumplir con este papel dentro de la familia, como se puede evidenciar en la figura 17, tanto madre como padre comparten el mismo porcentaje: 40.4%, el 17.8% tiene un jefe de familia diferente a padre o madre y existe el caso también en el que el estudiante es el jefe de familia con un 1.5% (Ver anexo 15).

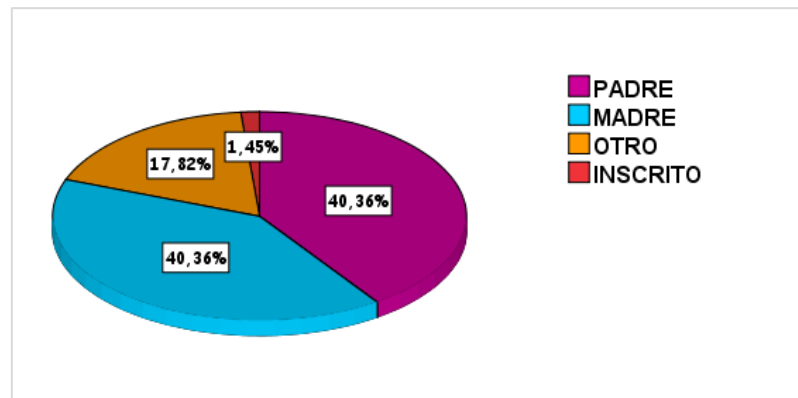


Figura 17. Distribución por jefe de familia.

Fuente: Esta investigación

**Distribución por ingresos familiares anuales:** Por ser una variable cuantitativa se ha agrupado para interpretar de mejor manera los resultados, hace referencia a aquellos ingresos económicos con los que cuenta una familia para cubrir sus necesidades básicas durante un año. Es un indicador muy importante porque refleja la calidad de vida que tiene el estudiante, lo cual puede repercutir en la facilidad o dificultad para permanecer en la carrera. Al analizar las cifras de la figura 18 se observa que más de la mitad de los estudiantes; el 54.9% tienen ingresos

familiares anuales entre cinco y diez millones de pesos, un valor promedio que se podría asociar al salario mínimo. Las familias del 28% de la población generan ingresos menores a cinco millones, siendo los hogares con más dificultades económicas. El 10.5% de los estudiantes tienen ingresos entre diez y quince millones, las familias de los estudiantes con más ingresos son muy pocas y comparten el mismo porcentaje, 3.2%. Según la tabla No. 7, el promedio de ingresos anuales es de \$7.926.051, con una desviación estándar de \$7.109.437, esta desviación indica que hay una amplia dispersión entre los diferentes valores, prueba de ello es el valor mínimo \$720.000 y el monto máximo de ingresos que es \$77.320.320 (Ver anexo 16).

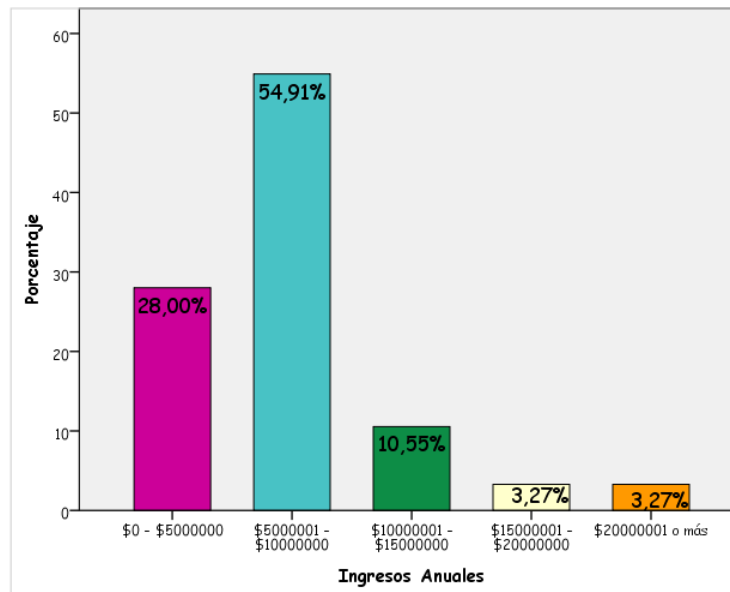


Figura 18. Distribución por ingresos familiares anuales.

Fuente: Esta investigación

**Distribución por puntaje socioeconómico:** Este factor se relaciona a la clasificación socioeconómica del estudiante, la cual es realizada por la universidad a partir de la documentación que se exige al momento de ingresar al programa, se hace con el fin de asignar el valor de los derechos de matrícula. Para una mejor interpretación de la información fue distribuida en rangos, de la figura 19 se puede inferir que el 61.1% de los estudiantes tiene un puntaje socioeconómico entre 6 y 10 puntos, los que están clasificados entre 11 y 15 puntos alcanzan un 28%, el 6.9% se ubican entre 16 y 20 puntos, por último los que tienen más de 20 puntos representan un 4% de la población. De acuerdo con la tabla No. 7, el promedio en el puntaje socioeconómico es de 11.19, con una desviación estándar de 4.64, un valor mínimo de 6 y un máximo de 40 puntos (Ver anexo 17).

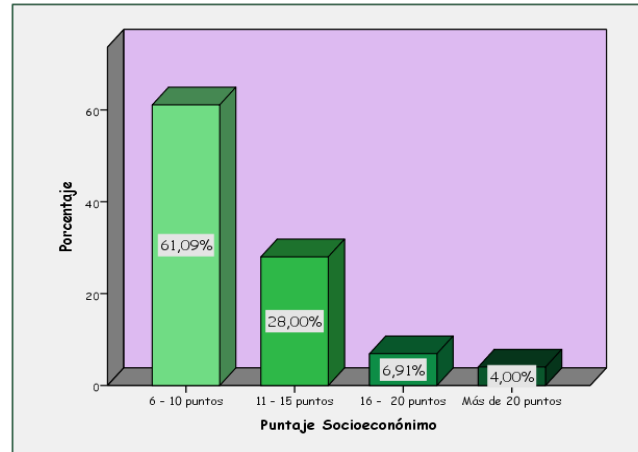


Figura 19. Distribución por puntaje socioeconómico.  
Fuente: Esta investigación

**Distribución por valor de matrícula:** El valor de la matrícula es una variable que hace referencia al costo en pesos que paga el estudiante cuando ingresa al primer semestre, este valor está estrechamente relacionado con el puntaje socioeconómico que la universidad le ha asignado. Según la figura 20 tan solo el 9% de los estudiantes pagan la matrícula más económica que está entre cero y cien mil pesos, más de la mitad de los estudiantes: 56% pagan entre cien mil y doscientos mil pesos, el 25% cancela un monto entre doscientos mil y trescientos mil, el resto de la población cancela matrículas superiores a los trescientos mil pesos. Según la tabla No.7, el promedio del valor de la matrícula es de \$210.363, con una desviación estándar de \$127.536, un valor mínimo de \$5.817 y un máximo de \$1.160.315 pesos (Ver anexo 18).

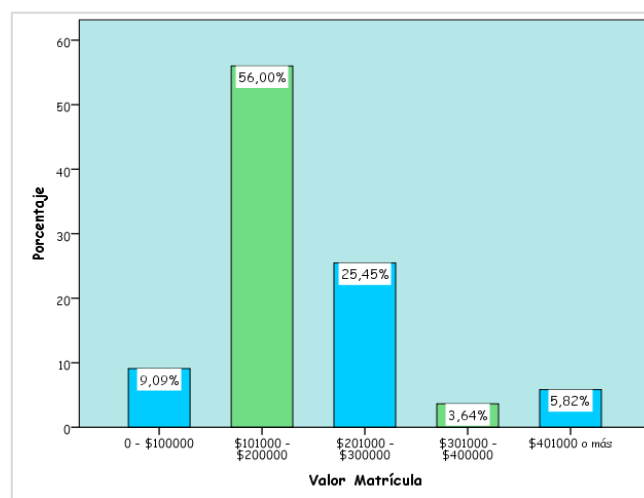
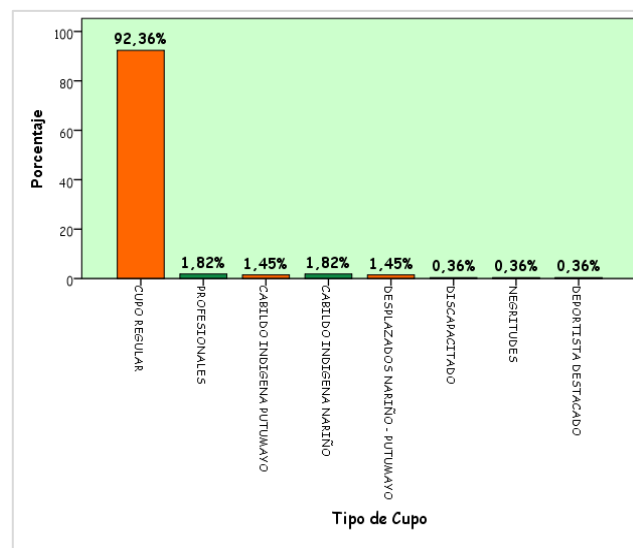


Figura 20. Distribución por valor de matrícula.  
Fuente: Esta investigación

### *Factores Institucionales*

**Distribución por tipo de cupo:** Dentro de los factores institucionales encontramos el tipo de cupo al momento de ingresar a la Universidad, el cual para la gran mayoría de estudiantes objeto de estudio: según la figura 21 el 92.3% fue regular, este cupo hace referencia a los estudiantes que no ostentan ninguna condición social o académica especial. El 7.7% restante de la población está distribuida en los diferentes tipos de cupo especiales; que son los que tienen una condición particular y para los cuales la universidad ofrece un tratamiento especial, como: profesionales, desplazados, cabildos indígenas, negritudes, etc. (Ver anexo 19).



*Figura 21.* Distribución por tipo de cupo.  
Fuente: Esta investigación

**Distribución por sanciones:** La sanción es una variable que se debe tener en cuenta ya que considera el régimen disciplinario y académico de la Universidad de Nariño, se sanciona a los estudiantes cuando cometen alguna falta leve, grave o gravísima o pierda materias de acuerdo con el estatuto estudiantil. Según la figura 22 la mayoría de los estudiantes del programa no han tenido ningún tipo de sanción con un porcentaje de 74.1%, el 13.4% de los estudiantes han tenido un promedio menor a 3.0 en el primer semestre. Han sido sancionados con pérdida temporal de derecho a continuar estudios un 9.4%, el resto de estudiantes está distribuido en los otros tipos de sanciones (Ver anexo 20).

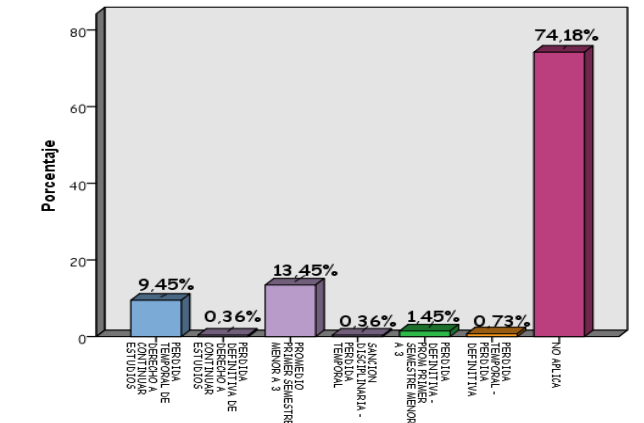


Figura 22. Distribución por sanciones.  
Fuente: Esta investigación

### *Deserción estudiantil*

La figura No. 23 muestra el consolidado en cuanto al porcentaje de estudiantes que desertaron del programa de Licenciatura en Informática en la cohorte establecida entre 2012-1 a 2016-2, en la cual se observa que un 41,45% de la población objeto de estudio finalmente abandonó la carrera, frente al 58,55% de los estudiantes que continúan con sus estudios (Ver anexo 21).

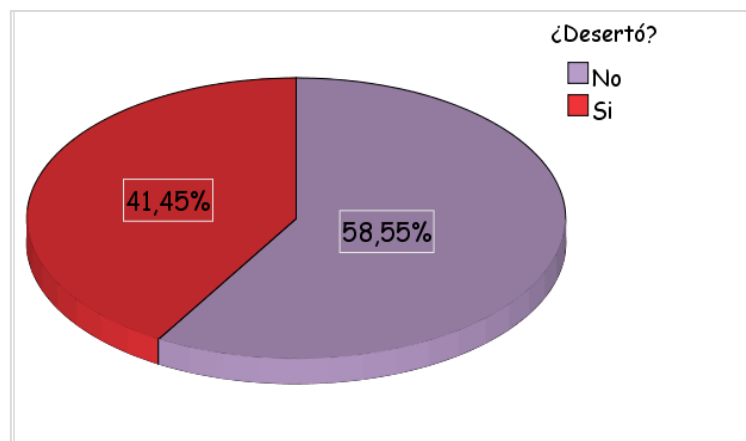


Figura 23. Deserción estudiantil en el programa de licenciatura en informática.  
Fuente: Esta investigación

Tabla 7

*Medidas de Tendencia Central y dispersión para variables cuantitativas*

Variable	Promedio	Desviación Estándar	Mínimo	Máximo
Edad Ingreso	19.70	4.06	16	32
Puntaje Icfes	56.7	4.55	42	77.35
Promedio de Notas	3.2	0.89	0	4.65
No. Materias perdidas en total	3.6	2.9	0	17
Ingresos familiares anuales	7926051	7109437	720000	77320320
Puntaje socioeconómico	11.19	4.64	6	40
Valor de la matrícula	210363	127536	5817	1160315

Fuente: Esta investigación

#### 4.2.2 Análisis Bivariado.

Con el fin de establecer la relación de los factores asociados al fenómeno de la deserción y conocer su papel predictor, se realiza un análisis bivariado; el cual se constituye en una técnica que permite encontrar el grado de relación entre las distintas variables independientes frente a la variable dependiente, en este caso si el estudiante deserto o no del programa.

Según la revisión de la literatura, lo más recomendable es explorar las asociaciones bivariantes mediante un análisis de tablas de contingencia, las cuales permiten representar simultáneamente dos variables para estudiar su relación. Para determinar la independencia se tendrá en cuenta el coeficiente Chi-cuadrado (*p-valor*) el cual indica el grado de confianza en la asociación de las variables, esta prueba asume una hipótesis inicial nula de que las variables no están relacionadas, por el contrario la hipótesis alterna señala dependencia entre las variables. Estos supuestos se contrastan con el nivel de significancia estadística que generalmente es menor a 0.05, este criterio es muy restrictivo y no permite incluir en el modelo variables que pueden ser fuertes predictoras cuando se incluyen con otras variables, por lo tanto para efectos de este



estudio; si el valor de la significación es menor o igual que 0.25 se admite la variable como factor predictivo, es decir, se acepta la hipótesis alterna.

Es así como obtenemos una primera aproximación a la medida de asociación, la OR. Para Cerda, Vera, & Rada (2013) “Matemáticamente un OR corresponde a un cociente entre dos odds, siendo un odds una forma alternativa de expresar la posibilidad de ocurrencia de un evento de interés o de presencia de una exposición” (pág. 1). Cuando su valor es 1 indica ausencia de asociación entre las variables, cuando es menor a 1 señala una asociación negativa y si su valor es mayor a 1 indica asociación positiva. El uso de OR se asocia a estudios donde se utilizan modelos de regresión logística como estrategia estadística para reducir el sesgo de confusión.

Este análisis se llevó a cabo en el programa estadístico SPSS versión 24, a continuación se presentan los resultados obtenidos.

#### ***4.2.2.1 Variables Cuantitativas.***

En el caso de las variables cuantitativas; se hace un análisis por separado y se obtiene el p-valor, para el cual se tendrá el mismo criterio que en las variables cualitativas, es decir las que superen el 0.25 no se tendrán en cuenta debido a la poca relación que tienen con la variable deserción.

**Tabla 8**

***Asociación de las variables cuantitativas con la variable deserción***

<b>Grupo</b>	<b>Variable</b>	<b>p-valor</b>
<b>Factores individuales</b>	Edad ingreso	0.995
<b>Factores académicos</b>	Puntaje Icfes	0.458
	Promedio de Notas	0.003*
	Último semestre cursado	0.000*
	No. de Materias perdidas último semestre	0.000*

	No. de Materias perdidas en total	0.000*
<b>Factores socioeconómicos</b>	Ingresos anuales familiares	0.726
	Puntaje socioeconómico	0.442
	Valor matrícula	0.008*

Fuente: Esta investigación

De la tabla anterior se puede determinar que las variables: edad de ingreso, puntaje Icfes, Ingresos anuales familiares y puntaje socioeconómico no son determinantes en el riesgo de deserción debido a que su significancia es mayor a 0,25. Por el contrario, las variables cuyo valor estadístico es significativo son: Promedio de notas, último semestre cursado, Número de materias perdidas en el último semestre, Número de materias perdidas en total y valor de matrícula.

#### 4.2.2.2 Variables Cualitativas

En el caso de las variables cualitativas, se hace uso de las tablas de contingencia a fin de obtener el p-valor. Para determinar el OR o riesgo se utilizó la regresión logística con cada variable, la cual toma como referencia (1) la última categoría y la compara con las demás categorías.

**Tabla 9**

*Asociación de factores individuales con la variable deserción*

VARIABLE		Si desertó	No desertó	OR	p-valor
<b>Sexo</b>	Femenino	32	55	0.75	0.285
	Masculino	82	106	1	

Fuente: Esta investigación

De la anterior tabla podemos inferir que no se encuentra relación entre la variable sexo y deserción ya que el valor de chi cuadrado no es significativo (p-valor= 0.285).

**Tabla 10***Asociación de factores académicos con la variable deserción*

VARIABLE		Si desertó	No desertó	OR	p-valor
<b>Tipo de colegio</b>	Oficial	96	151	1.27	0.018*
	Privado	17	8	4.25	
	Otro	1	2	1	

Fuente: Esta investigación

En cuanto a los factores académicos podemos afirmar que la variable Tipo de colegio resulta significativa (p-valor= 0.018), además una persona que viene de un colegio privado tiene 4 veces más riesgo de desertar que una de otro colegio.

**Tabla 11***Asociación de factores socioeconómicos con la variable deserción*

VARIABLE		Si desertó	No desertó	OR	p-valor
<b>Estrato Social</b>	1	50	84	0.32	0.074*
	2	50	70	0.39	
	3	13	7	1	
	4	1	0	0.00	
<b>Ciudad de Procedencia</b>	Pasto	80	119	0.59	0.585
	Municipio de Nariño	25	34	0.65	
	Otro	9	8	1	

<b>Lugar de Residencia</b>	Pasto urbano	85	115	1.53	0.433
	Pasto rural	16	19	1.74	
	Otro municipio	13	27	1	
<b>Tipo de Residencia</b>	Propia	64	114	0.61	0.090*
	Arrendada/Anticresada	36	32	1.23	
	Propia/paga cuotas	4	4	1.10	
	No informa	10	11	1	
<b>Vive con la Familia</b>	No	20	17	2.82	0.168*
	Si	89	132	1.61	
	No Informa	5	12	1	
<b>Tiene hermanos en la U</b>	No	103	131	2.14	0.039*
	Si	11	30	1	
<b>Jefe de Familia</b>	Padre	36	75	1.44	0.050*
	Madre	51	60	2.55	
	Otro	26	23	3.39	
	Inscrito	1	3	1	

Fuente: Esta investigación

Dentro de los factores socioeconómicos las variables que resultan significativas son: Estrato social, tipo de residencia, vive con la familia, tiene hermanos en la U y jefe de familia. Según el valor OR se puede interpretar que las personas que viven en arriendo o anticres tienen más riesgo de desertar, los estudiantes que no viven con sus familias tiene el doble de probabilidad al igual que los que no tienen hermanos en la universidad y para los que tienen como jefe de familia la categoría otro se triplica su probabilidad.

**Tabla 12***Asociación de factores institucionales con la variable deserción*

VARIABLE		Si desertó	No desertó	OR	p-valor
<b>Tipo de Cupo</b>	Cupo Regular	103	151	2.04	0.205*
	Profesionales	2	3	2	
	Cabildo indigena Putumayo	4	0	0	
	Cabildo indigena Nariño	2	3	2	
	Desplazados Nariño-Putumayo	1	3	1	
	Discapacitados	0	1	0	
	Negritudes	1	0	0	
	Deportista destacado	1	0	0	
<b>Sanciones</b>	Si	66	5	42.9	0.000*
	No	48	156	1	

Fuente: Esta investigación

En la tabla anterior observamos que las variables Tipo de cupo y sanciones tienen un p-valor menor a 0.25, por lo cual se afirma que son variables significativas. La variable sanciones se agrupó en dos categorías y de acuerdo con el OR se puede deducir que los estudiantes sancionados tienen 42 veces más probabilidad de desertar que las personas sin sanciones.

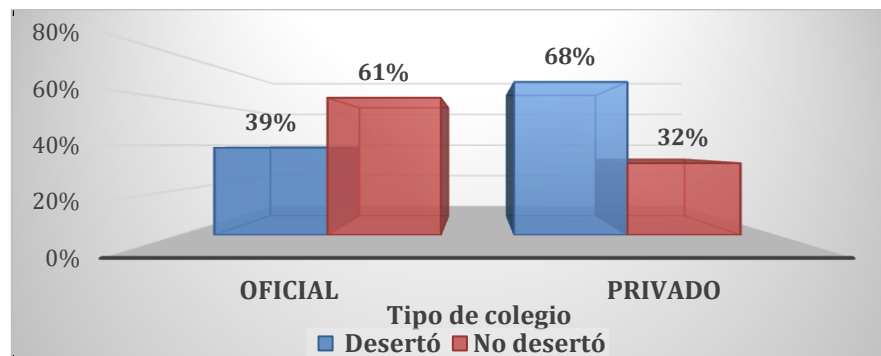
#### ***4.2.2.3 Análisis De Variables Significativas.***

Para determinar el perfil del estudiante desertor es importante analizar el comportamiento de las variables que resultaron significativas en el estudio, se determina el porcentaje de deserción de cada categoría dividiendo la frecuencia absoluta sobre el total. Para analizar las

variables cuantitativas se utiliza la agrupación en rangos, las frecuencias absolutas menores a 5 se obviaron para que la distribución sea adecuada y así evitar el sesgo, a continuación se muestran los resultados encontrados.

### *Variables Categóricas.*

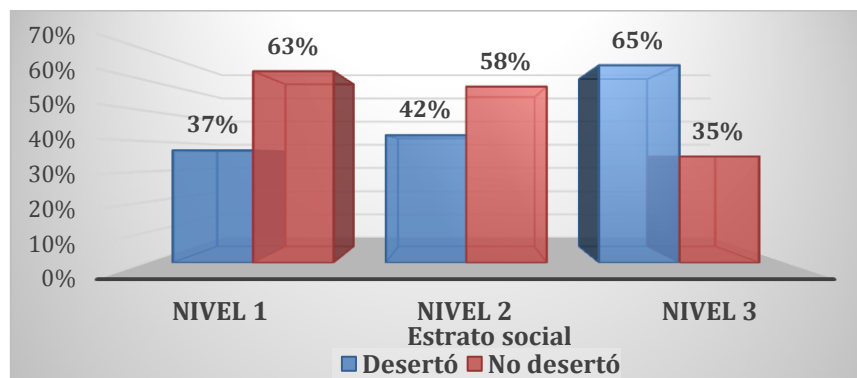
**Tipo de colegio:** En la siguiente figura se observa que el 68% de los estudiantes de colegio privado desertaron, por ende ellos son los que presentan mayor riesgo, en cuanto a los de colegio oficial el 39% han abandonado el programa.



*Figura 24.* Porcentaje de deserción según tipo de colegio.

Fuente: Esta investigación

**Estrato social:** De acuerdo con la figura 25, el mayor riesgo de deserción lo tienen los estudiantes de estrato 3 ya que de todos ellos desertó el 65%, del total de personas del nivel 2 desertó el 42% y de las personas de nivel 1 el 37%.



*Figura 25.* Porcentaje de deserción según estrato social.

Fuente: Esta investigación

**Tipo de residencia:** En la figura 26 se observa que el mayor porcentaje de deserción con un 53% se presentan en los estudiantes que viven en casa arrendada o anticresada, de todos los estudiantes con vivienda propia desertó el 36%, y las personas que no dieron información sobre esta variable desertaron el 48%.

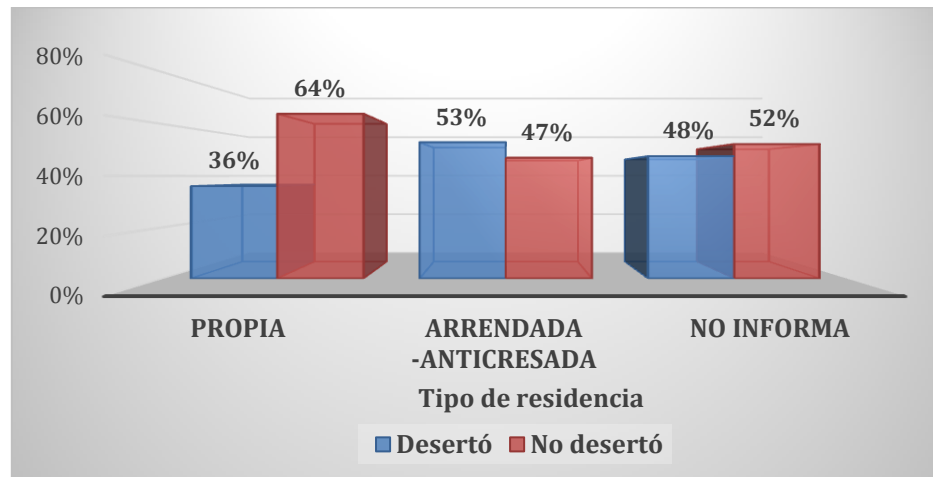


Figura 26. Porcentaje de deserción según tipo de residencia.  
Fuente: Esta investigación

**Vive con la familia:** En la siguiente figura se observa que el 54% de los estudiantes que no viven con la familia han desertado, por tanto son los que presentan mayor riesgo, de las personas que si viven con la familia el 40% han desistido de la carrera, así como el 29% de los que no informaron sobre esta variable.

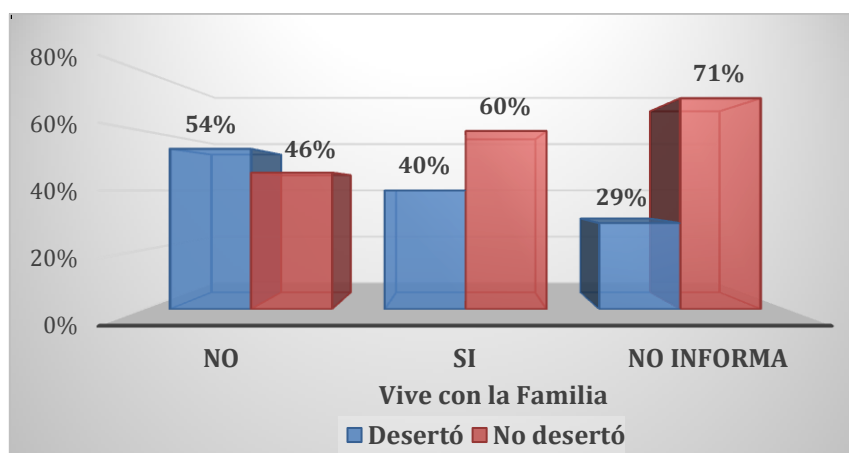
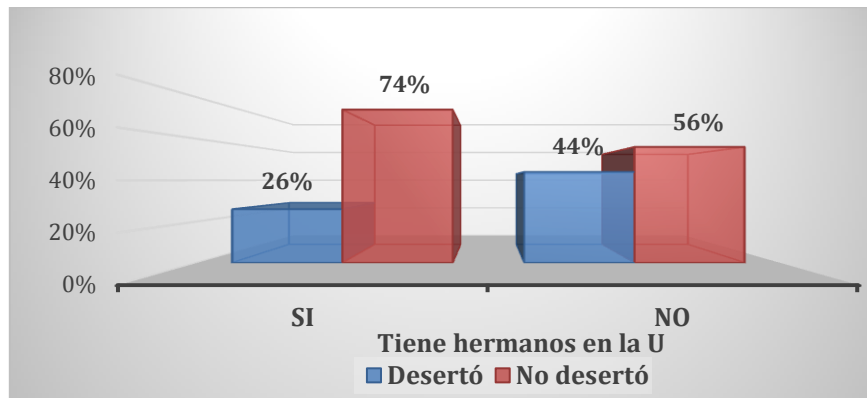


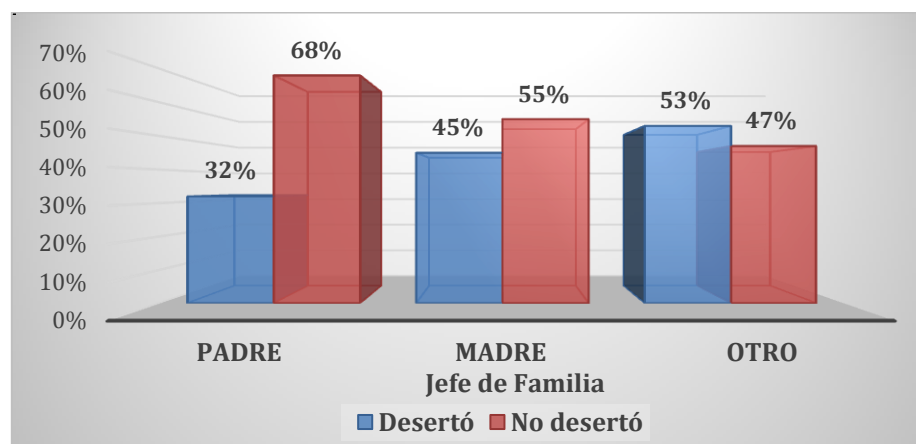
Figura 27. Porcentaje de deserción según vive con la familia.  
Fuente: Esta investigación

**Tiene hermanos en la U:** Según la figura 28, de todos los estudiantes que No tienen hermanos en la universidad desertaron el 44%, mientras que las personas que Si tienen hermanos dentro del plantel obtuvieron una menor proporción de deserción con un 26%.



*Figura 28.* Porcentaje de deserción según tiene hermanos en la U.  
Fuente: Esta investigación

**Jefe de familia:** En la figura 29 se observa que desertaron el 53% de los estudiantes que tienen como jefe de familia la opción otro, de todas las personas que tienen jefe de familia madre desertaron el 45%, y de las que tienen jefe de familia Padre el 32%.



*Figura 29.* Porcentaje de deserción según jefe de familia.  
Fuente: Esta investigación

**Tipo de cupo:** De acuerdo a la figura 30 se observa que el 59% de los estudiantes con otro tipo de cupo desertaron, por lo que tienen el mayor riesgo, el 41% de las personas de cupo regular abandonaron el programa.



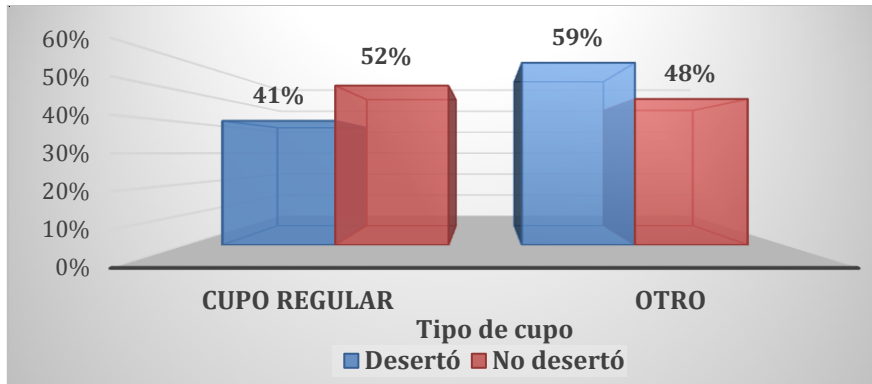


Figura 30. Porcentaje de deserción según tipo de cupo.  
Fuente: Esta investigación

**Sanciones:** De acuerdo con la siguiente figura los estudiantes que si tienen algún tipo de sanción por parte de la universidad tienen mayor riesgo de deserción ya que el 92% de ellos desertaron, de los que no tienen sanciones abandonaron el programa el 23%.

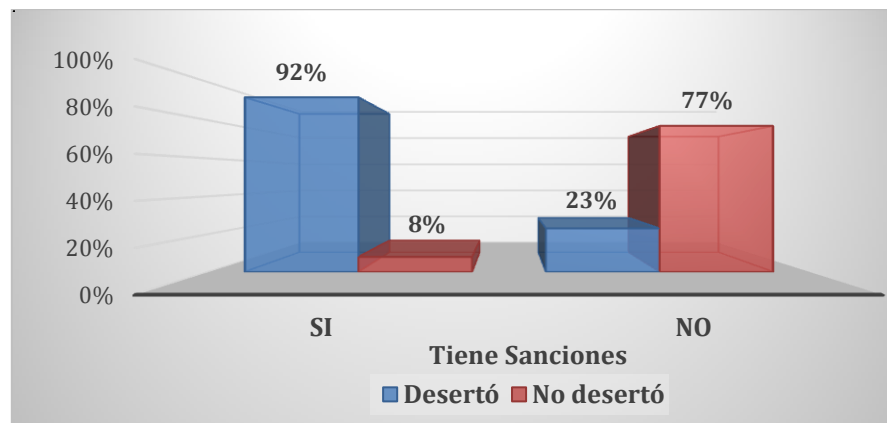


Figura 31. Porcentaje de deserción según sanciones.  
Fuente: Esta investigación

### **Variables Numéricas.**

**Valor de la matrícula:** En la figura 32 se puede observar que las personas con mayor riesgo de deserción de acuerdo al valor de la matrícula son las que pagan entre \$301.000 y \$400.000 con un porcentaje de deserción del 70%. Por el contrario los que menos porcentaje de deserción tienen con respecto a esta variable son los estudiantes que cancelan un valor entre 0 y \$100.000 pesos con un 12%.

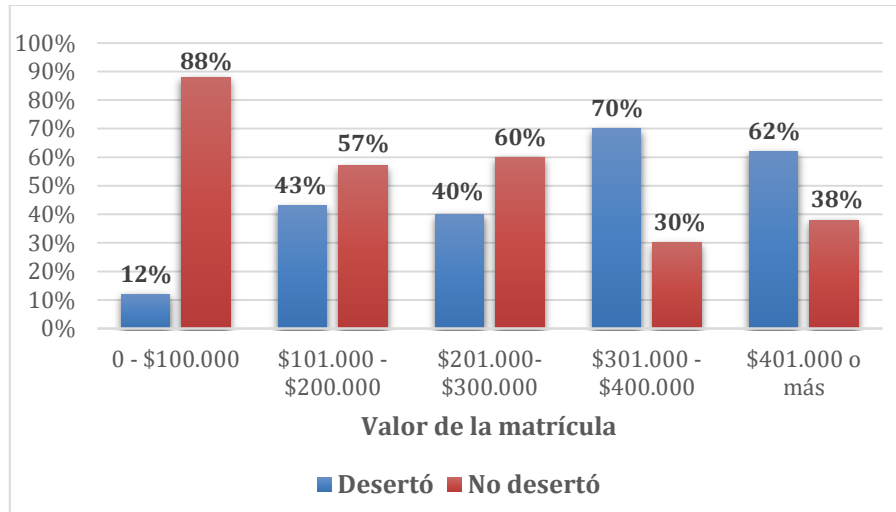


Figura 32. Porcentaje de deserción según el valor de matrícula.

Fuente: Esta investigación

**Promedio de Notas:** De acuerdo con la figura 33 los estudiantes que presentan mayor riesgo según el promedio de notas son los que tienen entre 0 y 2 puntos con un 100% de deserción, por el contrario los que menos riesgo tienen son los que sacan mayor promedio, entre 4.1 y 5 puntos.

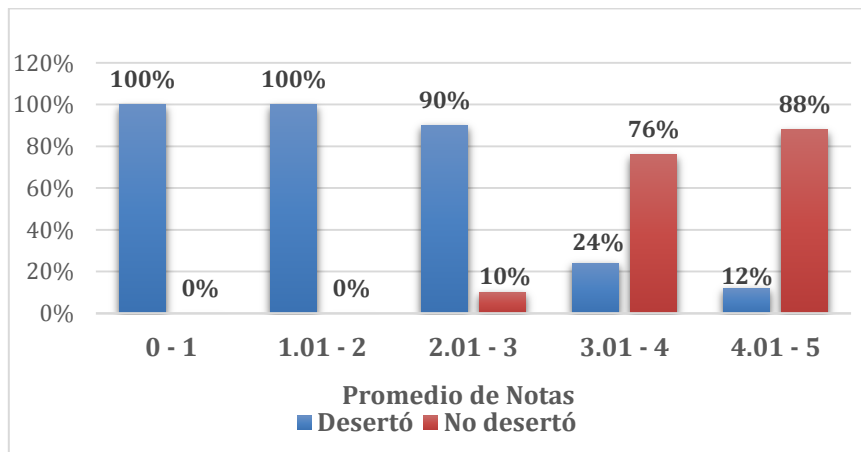


Figura 33. Porcentaje de deserción según promedio de notas.

Fuente: Esta investigación

**Último semestre cursado:** En la figura 34 se observa que las personas con mayor riesgo de desertar se ubican en los 2 primeros semestres de la carrera, mientras que los estudiantes de los últimos semestres tienen un porcentaje de deserción del 4%.

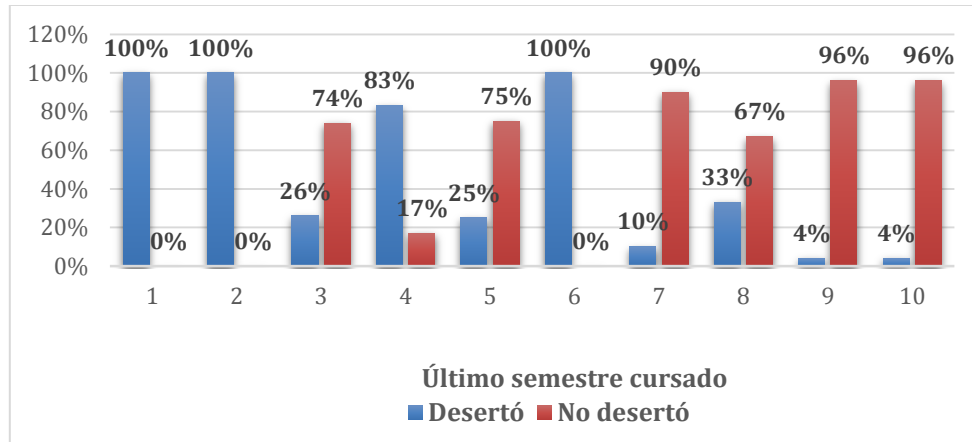


Figura 34. Porcentaje de deserción según el último semestre cursado.

Fuente: Esta investigación

**Número de materias perdidas en el último semestre:** En la siguiente figura se puede observar que los estudiantes que pierden más materias tienen un mayor riesgo de deserción, mientras que en los estudiantes que pierden menos materias aumenta su porcentaje de permanencia.

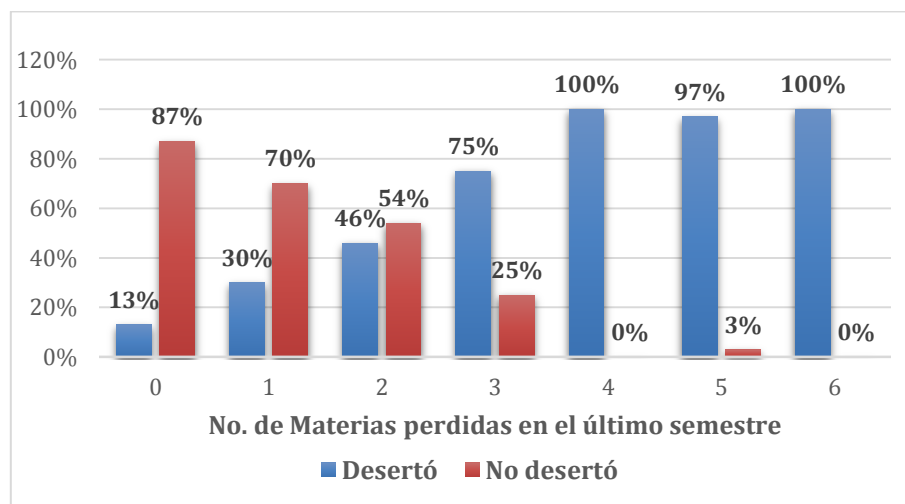


Figura 35. Porcentaje de deserción según materias perdidas en el último semestre.

Fuente: Esta investigación

**Número de materias perdidas en total:** Según la figura 36 el 86% de Los estudiantes que tienen 12 o más materias perdidas han desertado, siendo los que mayor riesgo presentan, mientras que de las personas que perdieron entre 0 y 3 materias tan solo el 23% desertó.

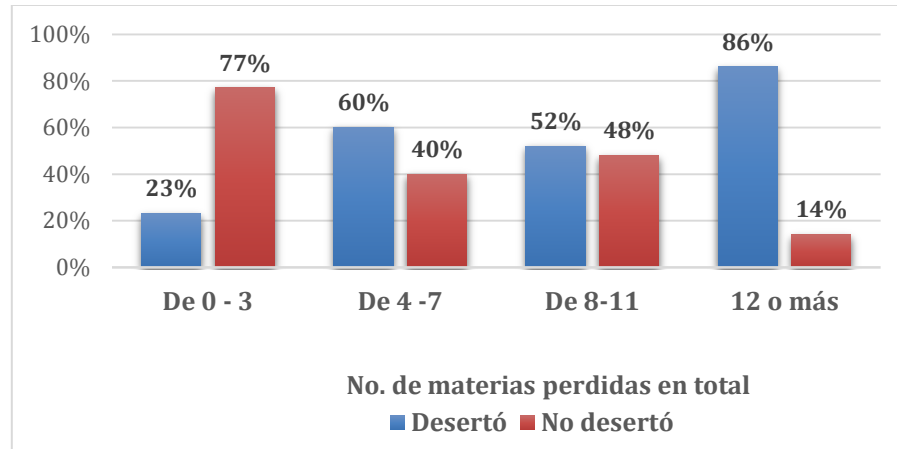


Figura 36. Porcentaje de deserción según el número de materias perdidas en total.  
Fuente: Esta investigación

#### 4.2.2.4 Perfil del Estudiante desertor.

De acuerdo con la información anterior se confirma que según las variables categóricas el estudiante desertor del programa es una persona que: proviene de un colegio privado, pertenece a un estrato social de nivel 3, su vivienda es arrendada o anticresada, no vive con su familia, no tiene hermanos en la universidad de Nariño, el jefe de familia en su hogar es una persona distinta al padre o la madre, tiene otro tipo de cupo y por lo general ha recibido alguna sanción por parte de la universidad. Según las variables numéricas el estudiante desertor paga una matrícula de \$400.000 pesos o más, tiene un promedio de notas inferior a 3, cursa en los dos semestres de la carrera, ha perdido más de 4 materias en el último semestre y en total a lo largo de la carrera ha perdido 12 o más materias.

#### 4.2.3 Análisis Multivariado.

Después de la revisión de la literatura existente se decide realizar un análisis multivariado a través del modelo de regresión logística binaria, para confirmar los factores observados que explican la presencia del fenómeno en cuestión. Este modelo es una técnica estadística muy útil cuando se tiene una variable dependiente (Deserción) y un conjunto de variables independientes, al construir el modelo se obtiene una ecuación que calcula o predice el riesgo en un nuevo individuo con cierta validez y precisión.

Para desarrollar el modelo predictivo es necesario conocer los factores que se asocian a la deserción, los cuales se obtuvieron en el análisis anterior. Se clasificó las variables según el valor en la significación estadística de Chi cuadrado ( $\leq 0.25$ ) obteniendo así 13 variables que guardan relación con la probabilidad de que un estudiante deserte.

#### 4.2.3.1 Codificación de variables.

Para simplificar la interpretación y análisis del modelo se debe identificar correctamente las categorías de las variables dicotómicas, en el caso de las variables numéricas no se hizo ningún cambio, las variables: Tipo de colegio, estrato social, tipo de residencia, vive con la familia, jefe de familia y tipo de cupo tienen más de dos categorías, por lo tanto se codifica usando variables indicadoras (*dummy*).

**Tabla 13**

#### *Codificación de la variable dependiente*

Variable Dependiente		
Deserción del estudiante	Categoría	Codificación
	No	0
	Si	1

Fuente: Esta investigación

**Tabla 14**

#### *Codificación de las variables independientes*

Variables Independientes		
Variable	Tipo de variable	P- valor
Tipo de colegio	Nominal Politómica	0.018
Estrato social	Nominal Politómica	0.074

Tipo de residencia	Nominal Politómica	0.090
Vive con la familia	Nominal Politómica	0.168
Tiene hermanos en la U---- No(0) Si(1)	Nominal Dicotómica	0.039
Jefe de Familia	Nominal Politómica	0.050
Tipo de cupo	Nominal Politómica	0.205
Sanciones-----No(0) Si(1)	Nominal Dicotómica	0.000
Valor matrícula	Numérica	0.008
Promedio de Notas	Numérica	0.003
Último semestre cursado	Numérica	0.000
No. de materias perdidas en el último semestre	Numérica	0.000
No. de materias perdidas en total	Numérica	0.000

Fuente: Esta investigación

#### ***4.2.3.2 Construcción del modelo de regresión logística binaria.***

Para la construcción del modelo se usó el software SPSS en su versión 24, empleando inicialmente un punto de corte de 0,5 (predeterminado del programa) esperando que este cumpliera con la predicción del riesgo de deserción, cabe destacar que el modelo permite obtener la probabilidad la cual es un valor que oscila entre 0 y 1, con este valor un estudiante puede ser clasificado en una de las dos categorías que tiene la variable dependiente, es decir si el valor de la probabilidad obtenida está entre 0 y 0,5 el estudiante es clasificado con 0 (No deserta), por el contrario si el cálculo es mayor a 0,5 se clasifica como 1 (si deserta).

Se empleó la técnica paso a paso de wald, la cual es un método automático por pasos hacia delante, que utiliza la prueba de Razón de Verosimilitud para comprobar las covariables a incluir o excluir en el modelo, en los resultados inicialmente aparece una tabla de clasificación con los valores observados y pronosticados.

Observado		Pronosticado		Porcentaje correcto	
		Deserción No desertó	Si desertó		
Paso 0	Deserción	No desertó	161	0	100,0
		Si desertó	114	0	,0
Porcentaje global					58,5

a. La constante se incluye en el modelo.  
b. El valor de corte es ,500

*Figura 37.* Clasificación inicial del modelo.  
Fuente: Esta investigación

Al analizar la figura 37 se observa que el grado de clasificación no es el esperado, el modelo tiene una especificidad alta, clasifica los verdaderos negativos en un 100% y una sensibilidad nula (0%) ya que no clasifica ningún verdadero positivo. Aunque codifica correctamente a los estudiantes que No desertan, pronostica mal a los que Si desertan, por tanto el modelo se limita debido a los falsos negativos que se obtendrían. Debido a esto fue necesaria la búsqueda de una estrategia que permita encontrar el punto de corte óptimo para que el modelo clasifique a los estudiantes de una mejor manera.

#### ***4.2.3.3 Elección del punto de corte.***

La metodología basada en distancias, propuesta por Costa, Boj del Val, & Fortiana (2012) permite encontrar el umbral óptimo a partir de una matriz de confusión en la que se calculan los verdaderos positivos (VP) y falsos positivos (FP); es decir, los VP son aquellos desertores pronosticados como desertores y los FP son desertores pronosticados como no desertores. En el caso de riesgo de deserción para determinar el punto de corte a partir del cual decidir si un estudiante deserta, se puede utilizar la denominada curva ROC la cual es medida importante que ayuda a evaluar el nivel de ajuste y clasificación del modelo, el programa SPSS nos permite realizar este procedimiento mostrando los siguientes resultados:

Resumen de procesamiento de casos	
Deserción	N válido (por lista)
Positivo <sup>a</sup>	114
Negativo	161

Figura 38. Procesamiento de casos.  
Fuente: Esta investigación

La muestra está compuesta por un total de 275 estudiantes de los cuales 114 desertaron y 161 no, la estimación de la curva ROC es:

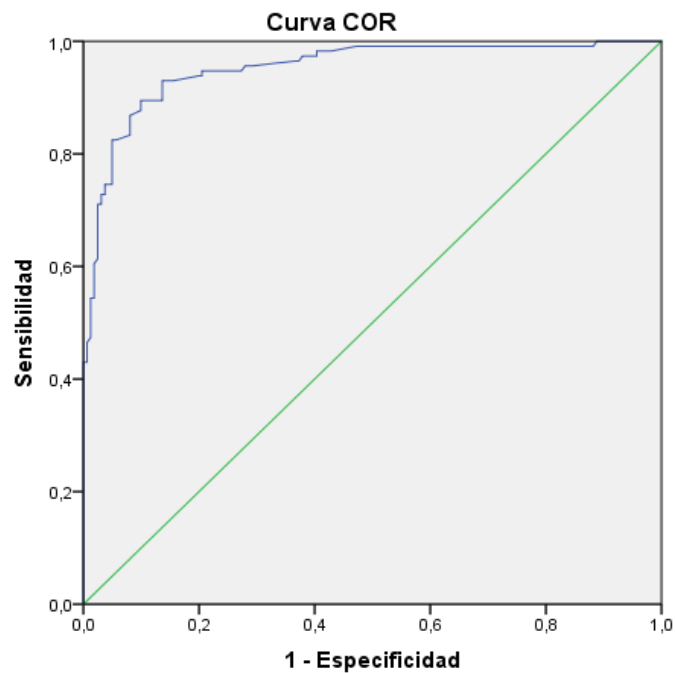


Figura 39. Curva ROC del modelo.  
Fuente: Esta investigación

En el eje X encontramos 1- Especificidad que es el porcentaje de FP y en el eje Y la Sensibilidad que representa los VP, se afirma que cuando la sensibilidad aumenta también lo hace el porcentaje de FP. Como se observa en la imagen el nivel de clasificación del modelo es bastante alto ya que la curva está alejada de la línea de no-discriminación.



**Área bajo la curva**

Variables de resultado de prueba: Probabilidad pronosticada

Área	Error estándar <sup>a</sup>	Significación asintótica <sup>b</sup>	95% de intervalo de confianza asintótico	
			Límite inferior	Límite superior
,952	,013	,000	,927	,977

*Figura 40. Área bajo la curva ROC.*  
Fuente: Esta investigación

El valor del área es de 95.2% con un intervalo de confianza de 95%, lo que permite afirmar que el modelo tendría una buena capacidad discriminante, el error estándar es 0.13, un valor bajo que indica la aproximación que tiene la curva al verdadero valor del área, la significancia es menor de 0.05 lo que señala que es altamente significativo.

De acuerdo con la matriz de confusión suministrada por el programa SPSS en la curva ROC se calcula el coeficiente Kolmogorov-Smirnov (K-S), el cual se basa en la función de distribución o probabilidades acumuladas para identificar el punto máximo de dicho coeficiente, el procedimiento es el siguiente:

1. Se ordena los diferentes puntos de corte de manera ascendente, los cuales están comprendidos entre 0 y 1.
2. Se estima la proporción acumulada de VP (sensibilidad) y FP (Especificidad) que comparten el mismo punto.
3. Se calcula las diferencias entre VP y FP por punto de corte.
4. Por último se identifica el punto de corte que tiene la máxima diferencia en el coeficiente K-S.

Debido a la magnitud de la tabla en la que se calculó el coeficiente K-S se muestra una extracción de ella:

Calculo del coeficiente K-S			
Punto de corte	Sensibilidad VP	1 - Especificidad FP	K-S
0,00	1,000	1,000	0,000
0,10	0,974	0,385	0,589
0,15	0,965	0,366	0,598
0,19	0,956	0,280	0,677
0,20	0,947	0,273	0,674
0,26	0,939	0,205	0,734
0,32	0,912	0,137	0,776
0,34	0,895	0,137	0,758
0,35	0,895	0,130	0,764
<b>0,39</b>	<b>0,895</b>	<b>0,099</b>	<b>0,795</b>
0,43	0,886	0,099	0,787
0,44	0,877	0,099	0,778
0,47	0,868	0,081	0,788
0,51	0,851	0,081	0,770
0,56	0,825	0,056	0,769
0,57	0,825	0,050	0,775
0,58	0,781	0,050	0,731
0,62	0,763	0,050	0,713

*Figura 41. Coeficiente K-S.*  
Fuente: Esta investigación

Al analizar la figura 41 podemos concluir que el punto exacto donde la sensibilidad es máxima y la especificidad es mínima es **0.39**, el cual ofrece un buen porcentaje de clasificación, diagnostica correctamente al 89.5% de verdaderos positivos y los falsos positivos tan solo alcanzan el 9.9%.

#### ***4.2.3.4 Aplicación del modelo de regresión logística binaria.***

Con la obtención del punto de corte óptimo se procede a realizar la regresión logística binaria utilizando la técnica paso a paso de wald empleada anteriormente, la cual retira aquellas variables que no tienen significancia estadística en el modelo. Se producen los siguientes resultados:

<b>Resumen del modelo</b>			
Paso	Logaritmo de la verosimilitud -2	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	237,276 <sup>a</sup>	0,390	0,525
2	183,158 <sup>b</sup>	0,499	0,672
3	164,735 <sup>b</sup>	0,531	0,716
4	158,000 <sup>b</sup>	0,543	0,731
5	152,691 <sup>c</sup>	0,551	0,743
6	153,688 <sup>c</sup>	0,550	0,740

Figura 42. Resumen del modelo.

Fuente: Esta investigación

En primer lugar aparece el resumen del modelo que permite evaluar su validez. Observamos que se registraron 6 pasos en total que indican los cambios que se han ido generando paso a paso. El valor de 2-LL (Logaritmo de la verosimilitud -2) mide el ajuste de los datos; los dos últimos pasos son los de menor valor, por lo que su precisión es mejor. El valor de R cuadrado de Cox y Snell y R cuadrado de Nagelkerke indican que el modelo explica el 74% de la varianza de la variable dependiente. Por lo cual podemos afirmar que las variables independientes incluidas en el modelo permiten predecir de una manera adecuada cuál puede ser resultado de la variable dependiente.

<b>Tabla de clasificación<sup>a</sup></b>					
Observado		Pronosticado			
		Deserción del Estudiante		Porcentaje correcto	
		No desertó	Si desertó		
Paso 6	Deserción	No desertó	145	16	90,1
		Si desertó	12	102	89,5
Porcentaje global					89,8

a. El valor de corte es ,395

Figura 43. Tabla Final de Clasificación.

Fuente: Esta investigación

Luego encontramos la tabla de clasificación que permite evaluar el ajuste del modelo comparando los valores pronosticados con los observados, de ella se concluye que el modelo tiene la capacidad de clasificar correctamente el 89.8% de los estudiantes analizados, obteniendo

una gran mejoría en cuanto a la clasificación. Los verdaderos negativos o No desertores han sido diagnosticados en un 90.1% y los verdaderos positivos o desertores en un 89.5%.

#### 4.2.3.4 Variables en la ecuación.

Variables en la ecuación							
	B	Error estándar	Wald	gl	Sig.	Exp(B)	
Paso 6 <sup>º</sup>	Ultimo Semestre Cursado	-1,023	0,153	44,804	1	0,000	0,360
	Total Materias Perdidas	0,387	0,099	15,185	1	0,000	1,472
	Sanciones Universidad	1,866	0,706	6,984	1	0,008	6,461
	Jefe de familia madre	1,299	0,501	6,715	1	0,010	3,664
	Constante	1,355	0,440	9,483	1	0,002	3,878

Figura 44. Variables de la ecuación.  
Fuente: Esta investigación

En la anterior figura se observan las 4 variables finales que deja la ecuación, en efecto todas mantienen una significación estadística ( $<0,05$ ), mejorando la predicción de la variable dependiente: Deserción, la técnica de pasos hacia adelante de Wald excluyó automáticamente aquellas variables que no fueron significativas. Los valores de OR o Exp(B) están ajustados en cada variable e indican la fuerza de asociación con Deserción controladas todas las demás variables incluidas en el modelo.

De acuerdo con el valor de OR se puede interpretar lo siguiente: a medida que aumenta el valor de *Último semestre cursado* disminuye la probabilidad de deserción en un 64%, ( $=0,36-1$ ) dicho de otra manera; entre más semestres cursados tenga el estudiante menor riesgo. Para el *Total de materias perdidas*; por cada materia que pierda el estudiante aumenta la probabilidad de deserción en un 47% ( $=1,47-1$ ), por ende entre más asignaturas pierda mayor riesgo. Un estudiante que tenga alguna *Sanción* por parte de la Universidad tiene aproximadamente 6.5 veces más probabilidad de desertar que aquellos que no han tenido sanción. Si el *Jefe de Familia* del estudiante es la madre la probabilidad de abandonar los estudios es 3,7 veces mayor que aquellos que en los cuales el padre es el jefe de hogar. Con estos datos se elabora la ecuación de Regresión Logística:

$$P(Y=1) = \frac{e^{(\beta_0 + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n)}}{1 + e^{(\beta_0 + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n)}}$$

De forma simplificada:

$$P(Y=1) = \frac{e^z}{1 + e^z} = \frac{1}{1 + e^{-z}}$$

Para nuestro caso sería:

$$P(Y=1) = \frac{e^{(1.355 - 1.023*(\text{ÚltimoS.}) + 0.387*(\text{TotalM.}) + 1.866*(\text{Sanciones}) + 1.299*(\text{JefeF.}))}}{1 + e^{(1.355 - 1.023*(\text{ÚltimoS.}) + 0.387*(\text{TotalM.}) + 1.866*(\text{Sanciones}) + 1.299*(\text{JefeF.}))}}$$

Donde  $P(Y = 1)$  es la probabilidad de que un individuo deserte ya que el valor “1” se asigna a los desertores. Un ejemplo para calcular la probabilidad de deserción partiendo de este modelo sería:

Un estudiante de Licenciatura en Informática quien terminó el 4 semestre, durante la carrera ha perdido en total 6 materias, no tiene ninguna sanción por parte del comité curricular y en su hogar el jefe de familia es la madre.

Recordemos que las variables Sanciones y Jefe de familia madre están codificadas como 0=No y 1=Si.

En primer lugar calculamos el valor de **Z**:

$$Z = (1.355 - 1.023*(4) + 0.387*(6) + 1.866*(0) + 1.299*(1))$$

Al realizar el cálculo se obtiene que **Z = 0.884**

Entonces:

$$P(Y=1) = \frac{e^{(0.884)}}{1 + e^{(0.884)}} = \frac{2.42056}{3.42056} = 0.70$$

La probabilidad de que el estudiante deserte es de 70%, debido a que esta persona supera el valor del punto de corte optimizado para el modelo que es **0.39** clasifica como “1” Si deserta.

#### ***4.2.3.5 Propiedades De Ajuste Y Clasificación Del Modelo.***

Existen determinados instrumentos que permiten validar el modelo probabilístico construido, mediante el análisis de las propiedades de ajuste y clasificación, algunos de ellos como el test de Kolmogorov-Smirnov, chi-cuadrado, curva ROC y la matriz de confusión del modelo o tabla de clasificación ya fueron abordados en este trabajo, con los cuales se logró establecer el ajuste del modelo. Otro parámetro como la prueba de ómnibus ayudan a identificar la bondad de ajuste para saber si las variables que se incluyeron mejoran la predicción de la variable dependiente.

#### ***Prueba de Ómnibus.***

<b>Prueba ómnibus de coeficientes de modelo</b>				
		<b>Chi-cuadrado</b>	<b>gl</b>	<b>Sig.</b>
<b>Paso 6<sup>a</sup></b>	<b>Paso</b>	<b>-0,997</b>	<b>1</b>	<b>0,318</b>
	<b>Bloque</b>	<b>219,471</b>	<b>4</b>	<b>0,000</b>
	<b>Modelo</b>	<b>219,471</b>	<b>4</b>	<b>0,000</b>

*Figura 45. Prueba Ómnibus.*

Fuente: Esta investigación

Como se observa en esta prueba, la significación estadística es 0,000 ( $<0,05$ ), lo que indica que en el paso 6 y con las variables introducidas se mejora el ajuste de forma significativa, es decir, estas variables sí pueden predecir la deserción.

### **4.3 Sistematización Del Modelo Estadístico**

Para el desarrollo de una aplicación web que calcule la probabilidad de deserción en los estudiantes del programa de licenciatura en informática, se usó la metodología IWeb. Después de haber caracterizado al estudiante desertor y obtenido las variables predictoras mediante la aplicación de la regresión logística, ésta metodología va a permitir construir una aplicación práctica, confiable y adaptable, empezando con la formulación de la necesidad que se va a

resolver para terminar con el diseño, producción, evaluación y pruebas para la entrega de la aplicación web.

#### **4.3.1 Etapa 1: Formulación.**

Para realizar una correcta formulación es necesario preguntarnos, entre otras cosas:

*¿Cuál es la principal necesidad para la aplicación web?*

La aplicación web permite dar a conocer a los estudiantes de licenciatura en informática de manera fácil, la probabilidad de riesgo de deserción que cada uno tiene, de acuerdo con los valores de las variables independientes de cada estudiante: Último semestre cursado, número de materias perdidas en total, si el jefe de familia es la madre y si el estudiante tiene alguna sanción por parte de la Universidad de Nariño; obtenidas al aplicar la regresión logística.

*¿Cuál es el objetivo principal que debe satisfacer la aplicación web?*

La aplicación web posibilita estimar el riesgo de deserción en los estudiantes por medio de los valores de las variables ingresados de acuerdo con un caso específico. A partir de ello conocer si tiene una alta o baja probabilidad de desertar.

*¿Quién la va a usar?*

La aplicación web será usada principalmente por estudiantes del programa de licenciatura en informática de la Universidad de Nariño, quienes de manera muy fácil acceden a la aplicación sin autenticación e ingresan sus datos para la estimación del riesgo de deserción según sea el caso.

#### **4.3.2 Etapa 2: Planificación.**

Para el desarrollo de la aplicación web se tienen en cuenta las siguientes consideraciones:

##### **4.3.2.1 Factibilidad del Sistema.**

Luego de precisar los requerimientos y necesidades para la aplicación web propuesta, se considera establecer los aspectos que garanticen que la aplicación puede desarrollarse en perfectas condiciones y además cumpla los objetivos y metas propuestos. La factibilidad de la aplicación se desarrolla desde los siguientes aspectos:

*Factibilidad Técnica:* Entre los recursos para el desarrollo de la aplicación están:

**Tabla 15**

***Recursos Hardware***

Recursos	Función	Características	Ubicación
Servidor	Alojar la aplicación web	Dell Power Edge R810. Memoria Ram: 32GB Procesador: Intel Xeon 32 Núcleos. Disco duro: 320GB	Centro de datos Universidad de Nariño - Bloque Tecnológico
Computador desarrolladores	Desarrollo de la aplicación web	Procesador Intel Core I5. Memoria Ram de 6Gb. Disco duro de 700Gb. Conectividad a Internet.	-
Dispositivos usuarios	Permitir el acceso de los estudiantes a la aplicación web	Equipos portátiles o de escritorio con conectividad a internet	-

Fuente: Esta investigación

**Tabla 16**

***Recursos Software***

Recursos	Función	Características
WAMP Server	Realizar prácticas de la aplicación web	Versión 2.5 Entorno de desarrollo web que permite tener un servidor local
Sublime Text	Editar código fuente de la aplicación web	Versión 2.0.2 Software licenciado editor de texto y editor de código fuente, escrito en



		C++ y Python para los plugins, permite trabajar un periodo de prueba ilimitado
--	--	--

Fuente: Esta investigación

*Factibilidad Económica:* Para los factores económicos, se utilizan recursos que tiene la Universidad de Nariño en cuanto a la infraestructura para el montaje de la aplicación web. El software utilizado para la creación de la aplicación es gratuito o tiene un periodo de prueba limitado, por tal motivo no fue necesario hacer gastos adicionales por licenciamiento.

*Factibilidad Operativa:* La aplicación web es utilizada cada vez que un estudiante quiera calcular el riesgo de deserción ingresando los datos que la aplicación le solicita. Los administrativos y docentes se encargaran de garantizar el uso eficiente y constante por parte de los estudiantes, para poner en marcha estrategias que permitan disminuir los porcentajes de deserción en el programa.

#### **4.3.3 Etapa 3: Análisis.**

Después de examinar el problema, se definen aspectos fundamentales como el contenido, función, comportamiento e interacción que tendrá la aplicación web para la creación de un buen diseño de la misma.

La aplicación web se basa en la sistematización de un modelo estadístico que explica el comportamiento de las variables de acuerdo con el fenómeno de la deserción, razón por la cual se realiza una investigación previa, iniciando con la selección de datos que los facilita OCARA (Oficina de Control de Admisiones y Registro Académico) para luego procesarlos con el programa estadístico IBM-SPSS versión 24; posteriormente se aplica el modelo de regresión logística de donde se obtienen las variables significativas al momento que un estudiante decide desertar y la fórmula estadística que calcula dicho porcentaje. Con las variables y la fórmula se construye la aplicación.

Debido a la sencillez de la aplicación web da como resultado la implementación de un módulo.

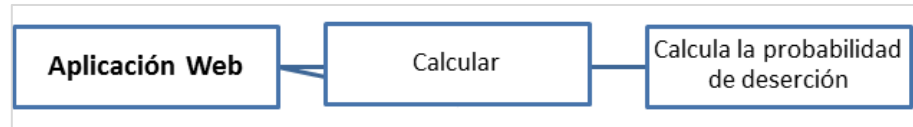


Figura 46. Módulo de la aplicación.

Fuente: Esta investigación

Para indicar la funcionalidad que ofrecerá la aplicación web es necesario mostrar la única categoría de usuario representada en el siguiente gráfico.

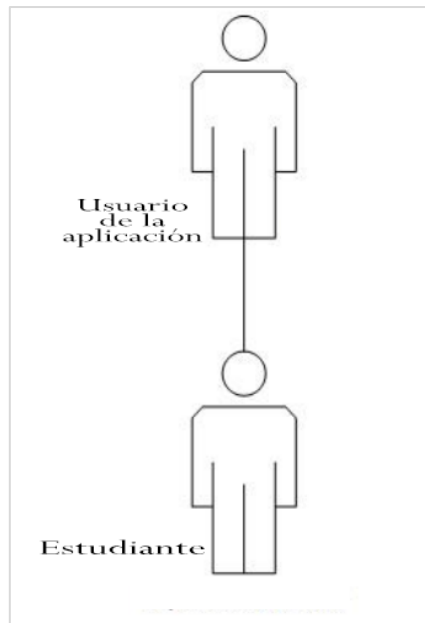


Figura 47. Usuario de la aplicación.

Fuente: Esta investigación

El usuario de la aplicación web puede realizar las siguientes actividades:

**Tabla 17**

***Rol y actividad que cumple el usuario***

Usuario	Rol	Descripción de actividades
Estudiantes	Visitante	Realizar el cálculo de la probabilidad de deserción

Fuente: Esta investigación

A continuación se presenta el caso de uso para el módulo de la aplicación para el usuario estudiante:

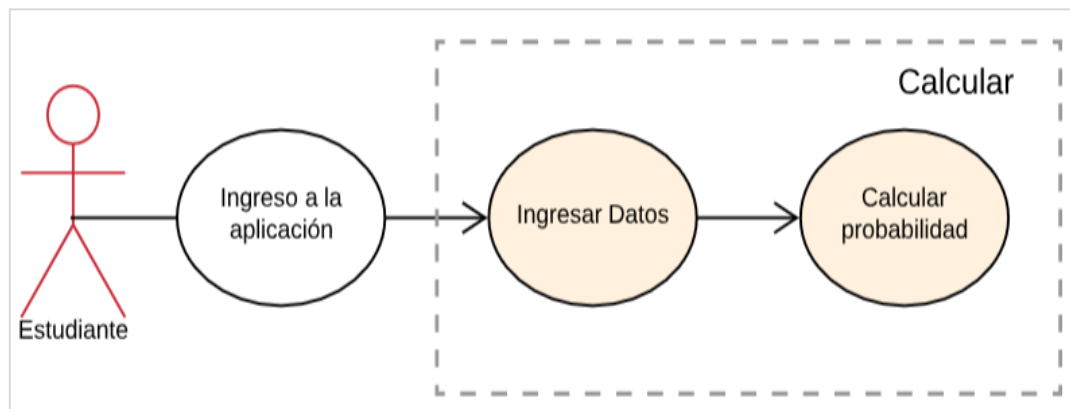
**Tabla 18**

*Casos de uso para el módulo del usuario estudiante.*

Caso de Uso	Descripción
Ingresar	Ingresar a la aplicación
Ingresar Datos	Escoger los valores de las variables
Calcular	Calcular la probabilidad de deserción

Fuente: Esta investigación

A continuación se muestran las actividades descritas anteriormente que el usuario Estudiante puede realizar en la aplicación web.



*Figura 48.* Diagrama de casos de uso de usuario Estudiante.

Fuente: Esta investigación

#### **4.3.4 Etapa 4: Ingeniería.**

En esta fase del proyecto se contemplan los siguientes aspectos: el diseño de navegación y el diseño de la interfaz gráfica.

#### 4.3.4.1 Diseño de Interfaz.

La interfaz gráfica para la aplicación web es muy sencilla ya que solo cuenta con el módulo de Calcular; se presenta en el siguiente esquema:

Zona 1: Header - Logo del programa de licenciatura en informática

Zona 2: Formulario - Espacios para ingresar los valores de las variables

Zona 3: Pie de página

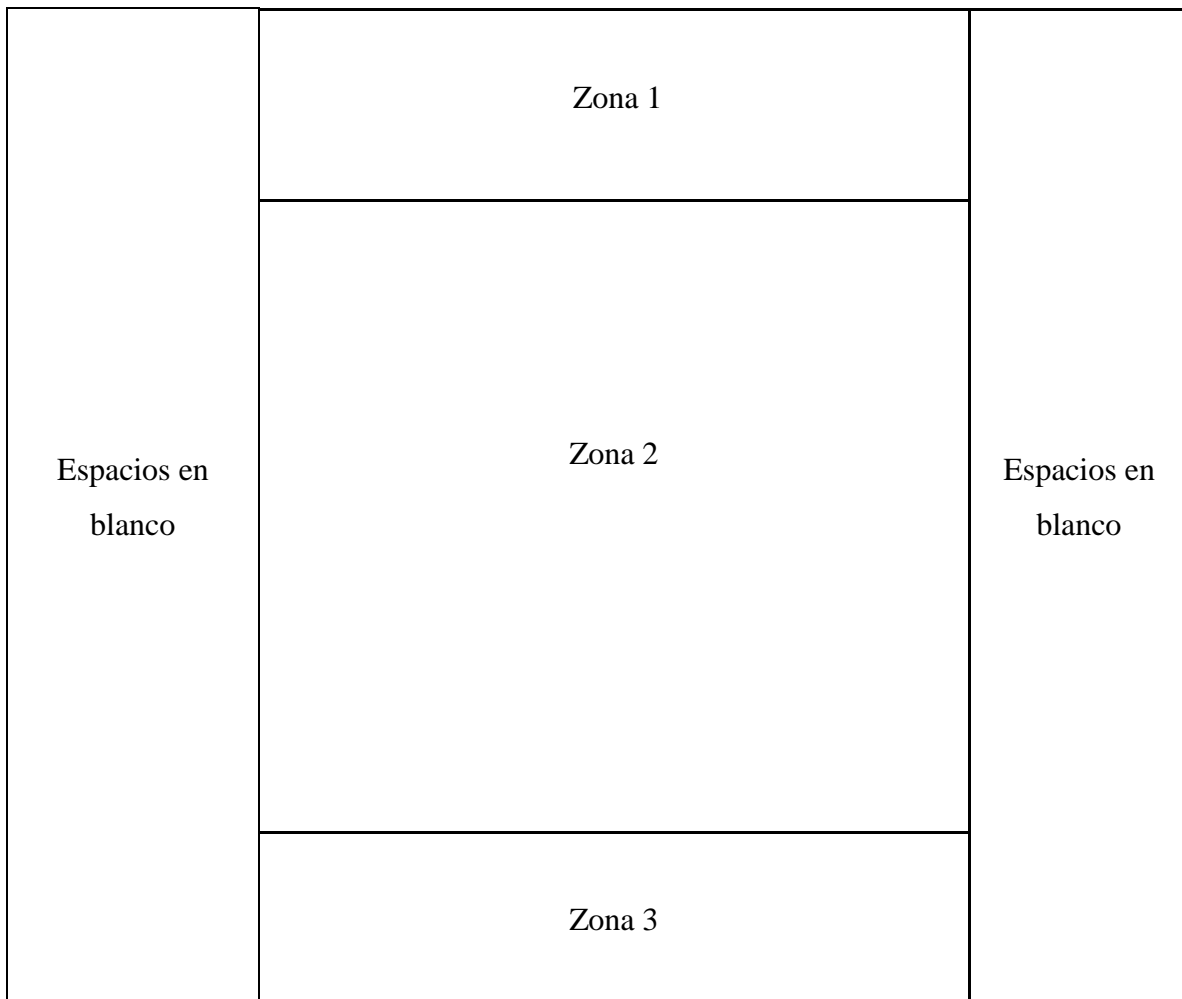


Figura 49. Diseño de Interfaz Gráfica.  
Fuente: Esta investigación

#### 4.3.4.2 Diseño de Navegación.

La ruta de navegación para el rol del usuario estudiante se realiza de la siguiente manera:

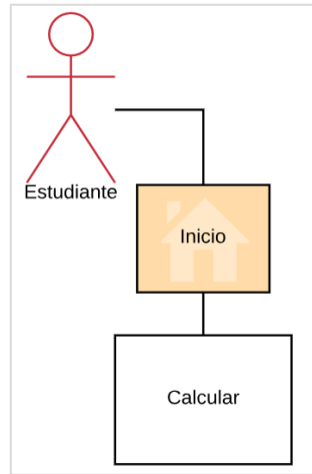


Figura 50. Diseño de Navegación.  
Fuente: Este estudio

#### ***Rol de actividades para el usuario Estudiante***

*Ingresar a la aplicación:* El usuario debe acceder a la aplicación ingresando a la dirección asignada para su uso en el servidor de la Universidad de Nariño.

*Calcular:* Después de haber ingresado el valor de las variables solicitadas en el formulario, el estudiante puede calcular la probabilidad de deserción y su resultado se muestra al pie del mismo formulario.

#### **4.3.5 Etapa 5: Generación de páginas.**

El entorno general en el cual fue desarrollada la aplicación web es la siguiente:

Lenguaje de programación: PHP 5.3.5, HTML5

Lenguaje para estilos: CSS

#### **4.3.6 Etapa 6: Pruebas y Evaluación.**

Para esta etapa se realizaron diferentes pruebas a la aplicación web en cada etapa anterior para la obtención de una aplicación web funcional. Se simuló estudiantes con diferentes datos para las variables solicitadas y se realiza el cálculo respectivo, con el fin de detectar posibles

errores en la fórmula se comprueba resolviendo la correspondiente fórmula en una calculadora y verificando la coincidencia de porcentajes con la aplicación web para cada uno de los estudiantes.

## 5. Estrategias

Un aspecto importante para el presente proyecto es la posibilidad de establecer estrategias que aporten a la retención estudiantil, a partir de la investigación que se realizó sobre la deserción. Como soluciones que contribuyen a la disminución de la deserción estudiantil se proponen las siguientes:

- Realizar actividades en las que se incluya el uso de la aplicación web desarrollada en el presente trabajo, con el fin de obtener el porcentaje de probabilidad individual para hacer un seguimiento a los estudiantes que se encuentren con mayor riesgo.
- Conformar un comité encargado de identificar y hacer seguimiento a los estudiantes que hayan perdido varias materias en el transcurso de la carrera, así como también a los que cuyo jefe de familia en el hogar es la madre y los que tienen alguna sanción por parte de la universidad, debido a que estos factores son los más determinantes en la deserción de acuerdo con la presente investigación.
- Gestionar y desarrollar actividades académicas de apoyo a fin de reforzar conocimientos habilidades y competencias que aporten a los estudiantes bases sólidas para enfrentar diferentes situaciones de la vida académica. Es importante dirigir estos esfuerzos a los estudiantes que se encuentren en los primeros semestres, los cuales según el desarrollo de esta investigación son los que presentan un mayor índice de deserción.
- Contribuir a la conformación y consolidación de grupos de estudio que apoyen el plan curricular y optimicen el rendimiento académico de los estudiantes, estos grupos pueden ser liderados por estudiantes con alto perfil académico que deseen colaborar en la superación de las dificultades que tienen sus compañeros.

- Fortalecer las actividades enfocadas a la orientación profesional y vocacional antes del ingreso al programa, la aplicación de pequeños talleres o tutorías dirigidos a estudiantes de secundaria puede considerarse como una opción. Dentro de estos talleres se puede abarcar temáticas que orienten al aspirante como: perfil profesional, competencias básicas, áreas, asignaturas, métodos de estudio, prácticas académicas, semilleros de investigación, etc.
- Consolidar, ampliar y reconocer la importancia de la información individual, académica, socioeconómica e institucional de cada estudiante que tiene la Universidad de Nariño en sus bases de datos, para permitir hacer un mejor seguimiento y control a los posibles desertores.

## Conclusiones y Recomendaciones

Después de todo el trabajo de investigación e interpretación de los resultados obtenidos de la aplicación del modelo estadístico y sistematización del mismo, se llegó a las siguientes conclusiones:

- El análisis realizado revela que la mayoría de variables que más influencia tienen a la hora de abandonar los estudios en la universidad son las que identifican al estudiante cuando ya ha ingresado al programa, como los factores académicos; a diferencia de factores individuales o socioeconómicos.
- El perfil del estudiante desertor se construyó gracias al análisis de los diferentes factores, los cuales dan una idea clara de las características específicas de un estudiante que posiblemente abandone la carrera. De acuerdo con esto, el perfil de un desertor corresponde a un estudiante que: proviene de un colegio privado, pertenece a un estrato social de nivel 3, su vivienda es arrendada o anticresada, no vive con su familia, no tiene hermanos en la universidad de Nariño, el jefe de familia en su hogar es una persona distinta al padre o la madre, tiene otro tipo de cupo, por lo general ha recibido alguna sanción por parte de la universidad, paga una matrícula de \$400.000 pesos o más, tiene un promedio de notas inferior a 3, cursa en los dos semestres de la carrera, ha perdido más de 4 materias en el último semestre y en total a lo largo de la carrera ha perdido 12 o más materias.
- El modelo de regresión logística finalmente arrojó cuatro variables predictoras que se asocian a la deserción las cuales son: Último semestre cursado, total materias perdidas, jefe de familia madre y sanciones. De ellas se puede inferir que; a medida que aumenta el último semestre cursado disminuye la probabilidad de deserción en un 64%, por cada materia que pierda el estudiante aumenta la posibilidad de desertar en un 47%, tener sanciones incrementa aproximadamente 6.5 veces más la probabilidad de desertar que aquellos que no han tenido sanción, si el Jefe de Familia del estudiante es la madre la probabilidad de abandonar los estudios es 3,7 veces mayor que aquellos que en los cuales el padre es el jefe de hogar.



- Al aplicar la técnica multivariante regresión logística, la cual es una herramienta estadística muy versátil, se observó que presenta los mejores resultados en la predicción de la deserción con las variables y categorías analizadas, puesto que tuvo la capacidad de clasificar correctamente el 89.8% de los estudiantes analizados, los verdaderos negativos o no desertores han sido diagnosticados en un 90.1% y los verdaderos positivos o desertores en un 89.5% siendo un modelo significativo y válido.
- Al sistematizar el modelo obtenido por medio de una aplicación web se logra identificar adecuadamente los estudiantes con posible deserción del programa de licenciatura en informática, siendo un recurso accesible y de fácil manejo que llame la atención a los individuos que quieran calcular su probabilidad de deserción.
- Los resultados obtenidos a lo largo de esta investigación permitieron proponer estrategias basadas en los factores que conllevan a que un estudiante deserte, estas estrategias aportan a la disminución de la deserción estudiantil y fortalecen los procesos de retención que realiza la Universidad de Nariño.
- El modelo propuesto por la presente investigación abre oportunidades para la creación de nuevos modelos de predicción utilizando técnicas estadísticas en los demás programas de la Universidad de Nariño que permitan un análisis comparativo de los factores que influyen en la deserción y el seguimiento constante a los estudiantes con alto riesgo, con el fin de reducir los efectos negativos dentro de la institución.

### Bibliografía

- Aguayo, C. (2007). Cómo hacer una Regresión Logística con SPSS© “paso a paso” . (I). *FABIS*, 1-16. Obtenido de [http://www.fabis.org/html/archivos/docuweb/Regres\\_log\\_1r.pdf](http://www.fabis.org/html/archivos/docuweb/Regres_log_1r.pdf)
- Alderete, A. (2006). Fundamentos del Análisis de Regresión Logística en la Investigación Psicológica. *Evaluar*, 6(1), 52-67. Obtenido de <https://revistas.unc.edu.ar/index.php/revaluar/article/view/534/474>
- Barón, F., & Téllez, F. (2004). *Apuntes de Bioestadística*. Universidad de Málaga. Obtenido de <http://www.bioestadistica.uma.es/baron/apuntes/ficheros/cap08.pdf>
- Berrocal, L. (2006). *Glosario Básico de términos estadísticos*. Lima, Perú: Instituto Nacional de Estadística e Informática. Obtenido de [https://www.inei.gob.pe/media/MenuRecursivo/publicaciones\\_digitales/Est/Lib0900/Libro.pdf](https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib0900/Libro.pdf)
- Borracci, R. (2001). El uso del riesgo relativo (RR), el odds ratio (OR) y la significación estadística "limítrofe". *Revista argentina de cardiología*, 69, 667-671. Obtenido de <https://www.sac.org.ar/wp-content/uploads/2014/07/1477.pdf>
- Buendía, L., Cólás, P., & Hernández, F. (2001). *Métodos de investigación en Psicopedagogía*. Madrid, España: McGraw-Hill Interamericana. Obtenido de [https://www.ugr.es/~ugr\\_unt/Material%20M%F3dulo%201/variables.pdf](https://www.ugr.es/~ugr_unt/Material%20M%F3dulo%201/variables.pdf)
- Calpa, F., & Cruz, D. (2008). *Análisis de la deserción estudiantil en la facultad de ciencias exactas y naturales de la universidad de nariño desde la cohorte 2001-2 hasta la cohorte 2006-2 utilizando el sistema spadies, (tesis de pregrado)*. Universidad de Nariño, Pasto.
- Casadesús, R. (2013). La estratificación social como la fuerza principal del cambio social. *ArsBrevis*(19), 56- 67. Obtenido de [goo.gl/Rfua5C](http://goo.gl/Rfua5C)
- Castañeda, M., Cabrera, A., Navarro, Y., & Vries, W. (2010). *Procesamiento de datos y análisis estadísticos utilizando SPSS*. Porto Alegre, Brasil: Edipucrs. Obtenido de <http://www.pucrs.br/edipucrs/spss.pdf>

- Castro, B., & Rivas, G. (2006). Estudio sobre el fenómeno de la deserción y retención escolar en localidades de alto riesgo. *Sociedad Hoy*, 11, 35-72. Obtenido de <http://www.redalyc.org/pdf/902/90201103.pdf>
- Cerda, J., Vera, C., & Rada, G. (2013). Odds ratio: aspectos teóricos y prácticos. *Rev Med Chile*, 141, 1329-1335. Obtenido de [https://scielo.conicyt.cl/scielo.php?script=sci\\_arttext&pid=S0034-98872013001000014](https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0034-98872013001000014)
- CNA. (2013). Lineamientos para la acreditación de programas de pregrado. 1-52. Obtenido de [https://www.cna.gov.co/1741/articles-186359\\_pregrado\\_2013.pdf](https://www.cna.gov.co/1741/articles-186359_pregrado_2013.pdf)
- Costa, T., Boj del Val, E., & Fortiana, J. (2012). Bondad de ajuste y elección del punto de corte en regresión logística basada en distancias. Aplicación al problema de credit scoring. 3<sup>a</sup> época(18), 19-40. Obtenido de [https://actuarios.org/wp-content/uploads/2017/02/anales2012\\_2.pdf](https://actuarios.org/wp-content/uploads/2017/02/anales2012_2.pdf)
- Cuadras, C. (2014). *Nuevos métodos de análisis multivariante*. Barcelona, España: CMCEditions. Obtenido de <http://www.ub.edu/stat/personal/cuadras/metodos.pdf>
- De la Fuente, S. (2011). *Análisis de variables categóricas. Tablas de contingencia*. Universidad Autónoma de Madrid, Madrid. Obtenido de <http://www.estadistica.net/ECONOMETRIA/CUALITATIVAS/CONTINGENCIA/tablas-contingencia.pdf>
- Del Valle, A. (s.f.). *Curvas ROC (Receiver-Operating-Characteristic) y sus aplicaciones*. Universidad de Sevilla, Sevilla. Obtenido de <https://idus.us.es/xmlui/bitstream/handle/11441/63201/Valle%20Benavides%20Ana%20Roc%20del%20TFG.pdf?sequence=1>
- Escobar, N. (2013). *Análisis de regresión logística para investigación de mercados*. Universidad Nacional de Colombia. Bogotá: Escuela de Administración y Contaduría Pública. Obtenido de <https://goo.gl/xyo2hd>
- Galeano, M. (2004). *Diseño de proyectos en la investigación cualitativa*. Medellín, Colombia: Fondo editorial Universidad EAFIT.

- Gómez, C., Padilla, A., & Rincón, C. (2016). Deserción escolar de adolescentes a partir de un estudio de corte transversal: Encuesta Nacional de Salud Mental Colombia. *Revista Colombiana de Psiquiatría*, 45(51), 105-112. Obtenido de <http://www.scielo.org.co/pdf/rcp/v45s1/v45s1a14.pdf>
- Granados, R. (2005). ¿Qué se entiende por sistematizar? *Enfermería en Costa Rica*, 26(1), 25-27. Obtenido de <http://www.binasss.sa.cr/revistas/enfermeria/v26n1/6.pdf>
- Hernández, R., Fernández, C., & Baptista, P. (2006). *Metodología de la investigación*. México D.F., México: McGraw Hill Interamericana. Obtenido de [https://www.esup.edu.pe/descargas/dep\\_investigacion/Metodologia%20de%20la%20investigaci%C3%B3n%20de%20Edici%C3%B3n.pdf](https://www.esup.edu.pe/descargas/dep_investigacion/Metodologia%20de%20la%20investigaci%C3%B3n%20de%20Edici%C3%B3n.pdf)
- Lazcano, E., Fernández, E., Salazar, E., & Hernández, M. (2000). Estudios de cohorte. Metodología, sesgos y aplicación. *Salud pública de México*, 42(3), 230-241. Obtenido de <http://www.scielosp.org/pdf/spm/v42n3/2858.pdf>
- Ley 115 Por la cual se expide la ley general de educación en Colombia (1994).
- Ley No. 30 Por el cual se organiza el servicio público de la Educación Superior en Colombia (1992).
- López, P., & Fachelli, S. (2016). *Metodología de la Investigación Social Cuantitativa*. Barcelona, España: Bellaterra (Cerdanyola del Vallès). Obtenido de [https://ddd.uab.cat/pub/caplli/2016/163570/metinvsoccua\\_a2016\\_cap3-10.pdf](https://ddd.uab.cat/pub/caplli/2016/163570/metinvsoccua_a2016_cap3-10.pdf)
- MEN. (2009). *Deserción estudiantil en la educación superior colombiana*. Bogotá. Obtenido de [http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articles-254702\\_libro\\_desercion.pdf](http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articles-254702_libro_desercion.pdf)
- MEN. (2012). *ACUERDO NACIONAL PARA DISMINUIR LA DESERCIÓN EN EDUCACIÓN SUPERIOR*. Bogotá. Obtenido de [http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articles-254702\\_archivo\\_pdf\\_politicas\\_estadisticas.pdf](http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articles-254702_archivo_pdf_politicas_estadisticas.pdf)

- MEN. (2015). *Guía para la implementación del modelo de gestión de permanencia y graduación estudiantil en instituciones de educación superior*. Bogotá. Obtenido de [http://www.mineduccion.gov.co/sistemasdeinformacion/1735/articles-254702\\_recurso\\_3.pdf](http://www.mineduccion.gov.co/sistemasdeinformacion/1735/articles-254702_recurso_3.pdf)
- MEN. (2016). *Boletín educación superior en cifras. Factores determinantes de la deserción*. Bogotá. Obtenido de [http://www.mineduccion.gov.co/1759/articles-358136\\_recurso.pdf](http://www.mineduccion.gov.co/1759/articles-358136_recurso.pdf)
- Murillo, C., & González, B. (2000). Manual de Econometría. Capítulo 6. 1-24. Obtenido de [http://www2.ulpgc.es/hege/almacen/download/7/7448/capitulo\\_6\\_dummy.PDF](http://www2.ulpgc.es/hege/almacen/download/7/7448/capitulo_6_dummy.PDF)
- Pardo, A., & Ruiz, M. (2005). *Análisis de datos con SPSS 13 Base*. Madrid, España: McGraw-Hill Interamericana.
- Pita, S., & Pértiga, S. (1997). Estadística descriptiva de los datos. *Jarpyo Editores*, 115-161. Obtenido de <https://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva2.pdf>
- Pressman, R. S. (2005). *Ingeniería del Software*. México: Mc Graww Hill.
- Saldaña, M. (2009). *Perfil del alumno desertor en la Universidad Católica de la Santísima Concepción. Un estudio de caso*. Universidad Católica de la Santísima Concepción-Chile, Concepción. Obtenido de Recuperado de: <https://goo.gl/Rr1FBN>
- Timarán, R., & Jiménez, J. (2015). Extracción de perfiles de deserción estudiantil en la Institución Universitaria CESMAG. *Ivestigium IRE: Ciencias Sociales y Humanas*, VI (1), 30-44.
- Tinto, V. (1989). Definir la deserción: una cuestión de perspectiva. *Revista de la Educación Superior*, 3(71), 1-9. Obtenido de [http://publicaciones.anuies.mx/pdfs/revista/Revista71\\_S1A3ES.pdf](http://publicaciones.anuies.mx/pdfs/revista/Revista71_S1A3ES.pdf)
- UDENAR. (2010-2015). *Udenar en cifras*. Pasto. Obtenido de <http://acreditacion.udenar.edu.co/wp-content/uploads/2015/11/0anuario.pdf>
- Vásquez, J., Castaño, E., Gallón, S., & Gómez, K. (2003). *Determinantes de la deserción estudiantil en la Universidad de Antioquia*. Medellín. Obtenido de <https://goo.gl/PkK1Tk>

**Anexos****Anexo 1. Distribución por Género.**

<b>Categoría</b>	<b>Frecuencia</b>	<b>%</b>
Femenino	87	31.6
Masculino	188	68.4

**Anexo 2. Distribución por Edad.**

<b>Rango</b>	<b>Frecuencia</b>	<b>%</b>
Menor de 20	181	65.8
20 - 24	73	26.5
25 - 29	19	6.9
30 o más	2	0.7

**Anexo 3. Distribución por Tipo De colegio.**

<b>Categoría</b>	<b>Frecuencia</b>	<b>%</b>
Oficial	247	89.8
Privado	25	9.1
No Informa	2	0.7
Valido Icfes	1	0.4

**Anexo 4. Distribución por Puntaje Icfes.**

<b>Rango</b>	<b>Frecuencia</b>	<b>%</b>
40 - 50.99	17	6.18
51 - 60.99	217	78.9
61 - 70.99	40	14.5
71 o más	1	0.3

**Anexo 5. Distribución por Promedio de notas.**

<b>Rango</b>	<b>Frecuencia</b>	<b>%</b>
0 - 1	12	4.3
1.01 - 2	15	5.4
2.01 - 3	50	18.1
3.01 - 4	155	56.3
4.01 - 5	43	15.6

**Anexo 6. Distribución por Último semestre cursado.**

<b>Semestre</b>	<b>Frecuencia</b>	<b>%</b>
1	57	20.7
2	16	5.8
3	65	23.6
4	6	2.2

5	40	14.5
6	2	0.7
7	39	14.2
8	3	1.1
9	24	8.7
10	23	8.4

**Anexo 7. Distribución por Materias perdidas en el último semestre.**

No.	Frecuencia	%
0	114	41.5
1	50	18.2
2	41	14.9
3	16	5.8
4	13	4.7
5	33	12
6	8	2.9

**Anexo 8. Distribución por Materias perdidas en total.**

Rango	Frecuencia	%
0 - 3	141	51.2
4 - 7	106	38.5



8 - 11	21	7.6
12 o más	7	2.5

**Anexo 9. Distribución por Estrato Social.**

Nivel	Frecuencia	%
1	134	48.7
2	120	43.6
3	20	7.3
4	1	0.36

**Anexo 10. Distribución por Lugar de procedencia.**

Categoría	Frecuencia	%
Pasto	199	72.4
Municipio de Nariño	58	21.5
Fuera de Nariño	17	6.1
Otro	1	0.3

**Anexo 11. Distribución por Lugar de residencia.**

Categoría	Frecuencia	%
Pasto urbano	200	72.7
Pasto Rural	35	12.7

Otro municipio	40	14.5
----------------	----	------

**Anexo 12. Distribución por Tipo de residencia.**

Categoría	Frecuencia	%
Propia	178	64.7
Arrendada/Anticresada	68	24.7
Propia/Pago cuotas	8	2.9
No Informa	21	7.6

**Anexo 13. Distribución por Vive con la familia.**

Categoría	Frecuencia	%
Si	221	80.4
No	37	13.5
No Informa	17	6.2

**Anexo 14. Distribución por Hermanos en la universidad.**

Categoría	Frecuencia	%
No	234	85.1
Si	41	14.9

**Anexo 15. Distribución por Jefe de familia.**

<b>Categoría</b>	<b>Frecuencia</b>	<b>%</b>
Padre	111	40.4
Madre	111	40.4
Otro	49	17.8
Inscrito	4	1.5

**Anexo 16. Distribución por Ingresos Anuales.**

<b>Rango</b>	<b>Frecuencia</b>	<b>%</b>
0 - \$5.000.000	77	28
\$5.000.001 - \$10.000.000	151	54.9
\$10.000.001 - \$15.000.000	29	10.5
\$15.000.001- \$20.000.000	9	3.3
\$20.000.001 o más	9	3.3

**Anexo 17. Distribución por Puntaje socioeconómico.**

<b>Rango</b>	<b>Frecuencia</b>	<b>%</b>
6-10 Puntos	168	61.1
11-15 Puntos	77	28.0
16 a 20 Puntos	19	6.9
Más de 20 Puntos	11	4

**Anexo 18. Distribución por Valor de matrícula**

Rango	Frecuencia	%
0 - \$100.000	25	9.1
\$101.000 - \$200.000	154	56
\$201.000- \$300.000	70	25.5
\$301.000 - \$400.000	10	3.6
\$401.000 o más	16	5.8

**Anexo 19. Distribución por Tipo de cupo.**

Tipo de cupo	Frecuencia	%
Cupo Regular	254	92,3
Profesionales	5	1,8
Cabildo Indígena Putumayo	4	1,4
Cabildo Indígena Nariño	5	1,8
Desplazados	4	1,4
Discapacitados	1	0,4
Negritudes	1	0,4
Deportista destacado	1	0,4

**Anexo 20. Distribución por Sanciones.**

Sanciones	Frecuencia	%
Pérdida temporal del derecho a continuar estudios	26	9.4
Pérdida definitiva del derecho a continuar estudios	1	0.4
Promedio Primer Semestre menor a 3	37	13.4
Sanción Disciplinaria- Pérdida Temporal	1	0.4
Pérdida definitiva- Promedio Primer Semestre menor a 3	4	1.4
Pérdida Temporal- Pérdida Definitiva	2	0.7
No aplica	205	74.1

**Anexo 21. Distribución por deserción estudiantil.**

Categoría	Frecuencia	%
Desertó	114	41.5
No desertó	161	58.5

**Anexo 22. Manual de Usuario**

**APLICACIÓN PARA EL CÁLCULO DE LA PROBABILIDAD DE DESERCIÓN**

**Por:**

**Michell Daniela Bravo**

**Haidy Tatiana Maya**

UNIVERSIDAD DE NARIÑO  
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA  
FACULTAD DE CIENCIAS EXACTAS Y NATURALES  
PROGRAMA DE LICENCIATURA EN INFORMÁTICA  
SAN JUAN DE PASTO  
2018

## Introducción

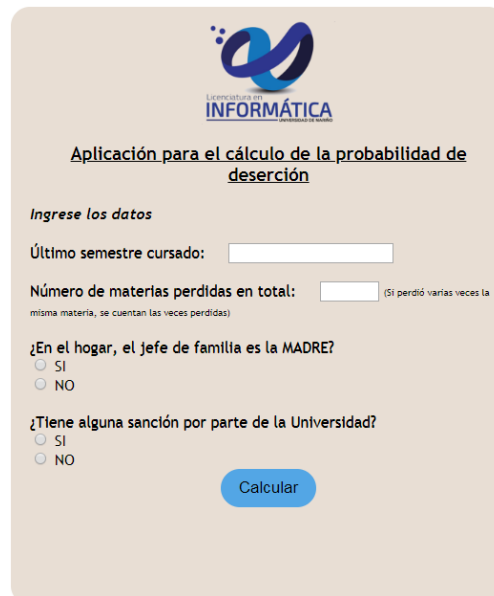
El presente manual es un instrumento útil que facilita, familiariza y es una guía para el uso de la aplicación web a las personas que sirvan como usuarios de la misma. Su contenido está relacionado con aspectos básicos del desarrollo de actividades de la aplicación.

### 1. ¿Qué es una aplicación web?

Una aplicación web es un programa informático donde el usuario accede desde un navegador web, sirve para realizar peticiones de tareas específicas y recibir una respuesta que se muestra en el mismo navegador.

### 2. Interfaz de la aplicación

En esta parte encontrar la interfaz de la aplicación, debido a su simplicidad solo cuenta con un esta única interfaz.



The screenshot shows a web application interface with a light beige background. At the top center is a logo consisting of a blue infinity symbol with the text 'Licenciatura en INFORMÁTICA' below it. Below the logo, the title 'Aplicación para el cálculo de la probabilidad de deserción' is displayed in bold. Underneath the title, the instruction 'Ingrese los datos' is shown. The form contains three input fields: 'Último semestre cursado:' with a text box, 'Número de materias perdidas en total:' with a text box and a note '(Si perdió varias veces la misma materia, se cuentan las veces perdidas)', and two radio button questions: '¿En el hogar, el jefe de familia es la MADRE?' and '¿Tiene alguna sanción por parte de la Universidad?'. At the bottom center, there is a blue 'Calcular' button.

### 3. Cálculo de la probabilidad de deserción

Una vez ingresada la dirección o url en la que se encuentra la aplicación, aparece su interfaz donde el estudiante debe seguir los siguientes pasos:

Paso 1: Escoja el valor del último semestre que cursó según sea el caso, entre primero y décimo semestre.

Paso 2: Ingrese el valor del número de materias perdidas en total, este valor también incrementa cuando el estudiante perdió varias veces la misma materia, por ejemplo, un estudiante de 5 semestre ha perdido dos veces matemáticas y una sola vez fundamentos de lógica, por consiguiente el dato que debe ingresar para esta variable es tres materias perdidas en total. Puede ingresar valores entre 0 y 20 materias.

Paso 3: Escoja entre las dos opciones (Si y No) si el jefe de familia en su hogar es la madre.

Paso 4: Escoja entre las dos opciones (Si y No) si tiene alguna sanción por parte de la universidad, durante su permanencia en el programa.

Paso 5: Presione el botón Calcular.

**Aplicación para el cálculo de la probabilidad de deserción**

*Ingrese los datos*

Último semestre cursado:

Número de materias perdidas en total:  (Si perdió varias veces la misma materia, se cuentan las veces perdidas)

¿En el hogar, el jefe de familia es la MADRE?

SI

NO

¿Tiene alguna sanción por parte de la Universidad?

SI

NO

**Calcular**

*Estudiante sus datos son:*  
4° Semestre, 6 Materias perdidas en total, El Jefe de Familia ES la madre y No tiene sanciones  
**Su probabilidad de deserción es 70.77 %**

En la parte de abajo de la aplicación puede observar los datos que previamente seleccionó y el resultado de la probabilidad que el mismo le genera.