

CAPITULO II

II. MARCO TEÓRICO

2.1 INVESTIGACIÓN DE MERCADO

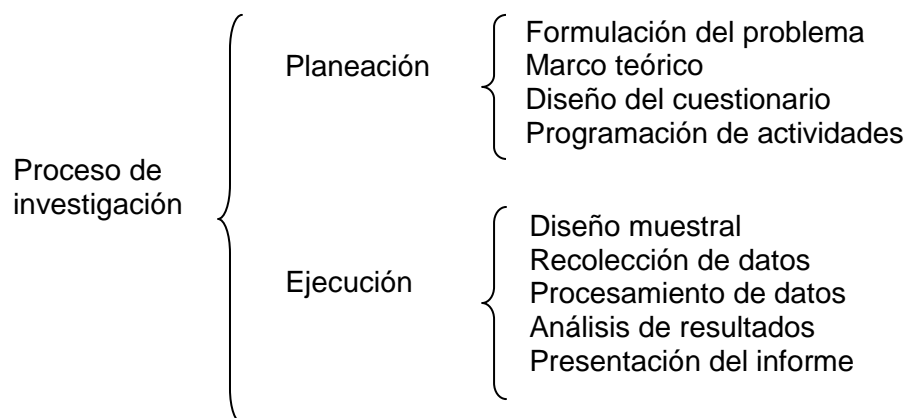
2.1.1 Concepto

La investigación de mercado es la técnica que trata de la recolección sistemática y objetiva de información, construcción de modelos y obtención de hallazgos y hechos que permitan tomar decisiones relacionados con problemas de mercado de bienes y servicios.

La investigación de mercado se origina ante la necesidad del gerente por minimizar los riesgos comerciales del bien o servicio que mercadea o piensa mercadear.

2.1.2 Proceso de investigación de mercado

El proceso de investigación de mercado comprende dos etapas básicas: la planificación y la ejecución. La planeación tiene por objetivo formular el problema, enfocar la investigación y diseñar las acciones a ejecutar. La segunda comprende la ejecución y el control de la investigación. El proceso de investigación tiene nueve fases, las cuales se presentan en el esquema siguiente:



2.1.3 Definición del problema

En la actualidad la mayoría de compañías en Guayaquil implementan sistemas informáticos para automatizar ciertos procesos bien sea para minimizar el tiempo de atención al cliente o para mantener informados al personal laboral, pero ciertas veces no basta con tener un ordenador conectado en red, porque si el cliente necesita una información necesariamente tendrá que acceder a una computadora esperando que no haya ningún error de conexión y que exista disponibilidad para acceder a la información. Es por esto que algunas compañías en la actualidad han optado por implementar terminales de información para diversas áreas como servicio al cliente, se conoce además que es utilizado como medio de publicidad o para el turismo; siendo estas muy poco explotadas dentro del sector empresarial en esta ciudad, tal vez esto se debe a la desinformación acerca de estos por lo que se necesita realizar una investigación de mercado para conocer que tan factible resultará

crear una empresa que se dedique a prestar los servicios de instalación, alquiler y mantenimiento de terminales de información.

2.1.4 Objetivos de la investigación

Los objetivos de la investigación son los siguientes:

- Determinar el conocimiento de los empresarios acerca de terminales de información.
- Conocer si las compañías están dispuestas a implementar en su negocio los terminales de información.
- Establecer en que áreas de la organización pretenden implementar los terminales de información.
- Determinar a que tipos de compañías le interesaría adquirir un Terminal, clasificándolas por actividad económica.
- Saber si desearían comprar o alquilar terminales de información.
- Conocer las características más relevantes que las compañías observan en una empresa que presta servicios referentes a terminales de información..
- Conocer si las empresas desearían participar como futuros clientes o como patrocinadores de SINTEC.
- Determinar los segmentos de mercado a los cuales se debe enfocar SINTEC.

2.1.5 Plan e implantación de la investigación

Los métodos de investigación son el plan básico que guiará la fase de investigación y análisis de datos del proyecto de terminales de información. Es la estructura que especificará el tipo de información requerida y sus fuentes de datos.

Para este caso de los terminales de información se utilizará la investigación descriptiva, puesto que esta tiene como propósito dominar o ampliar el conocimiento de la estructura, propiedades y medidas de los mercados. El instrumento de medición es la encuesta y es el prototipo que caracteriza la investigación de mercado, para esta herramienta la medición es lo fundamental, pero no se descarta los propósitos cualitativos.

2.2 Definiciones Básicas Estadísticas

Coefficiente de sesgo

Es una medida relativa que permite describir la asimetría de los datos alrededor de la media, tenemos tres casos: cuando el coeficiente de sesgo es negativo la mayor concentración de datos se encuentran hacia la derecha de la media, es decir que está sesgada hacia la izquierda; cuando el coeficiente de sesgo es positivo la mayor concentración de datos se encuentra hacia la izquierda de la media, es decir que está sesgada a la derecha; y cuando el coeficiente de sesgo es cero la media y la mediana son iguales, es decir que la distribución es simétrica. El coeficiente de sesgo es calculado de la siguiente manera:

$$\gamma_1 = \frac{\frac{1}{N} \sum_{i=1}^n (x_i - \bar{X})^3}{\left[\frac{1}{N} \sum_{i=1}^n (x_i - \bar{X})^2 \right]^{3/2}}$$

Coefficiente de kurtosis

Es una medida relativa, que permite establecer el grado de apuntamiento o achatamiento de la curva de distribución comparada con la distribución normal, al igual que en el sesgo, tenemos tres casos; distribución mesocúrtica cuando tiene la forma de una normal y su coeficiente es igual a tres, distribución platicúrtica cuando es achatada con respecto a una normal y su coeficiente es menor a tres; y por último, distribución leptocúrtica cuando es más apuntada que una normal y su coeficiente es mayor a tres; así mismo este coeficiente se calcula a través de la relación entre el cuarto momento central y la varianza al cuadrado como se muestra a continuación:

$$\alpha_4 = \frac{\mu_4}{\sigma^4} = \frac{E[(x_i - \mu_4)^4]}{\sigma^4}$$

Para determinar el estimado del coeficiente de kurtosis se tiene lo siguiente:

$$\bar{\alpha}_4 = \frac{\bar{\mu}_4}{S^4} = \frac{\frac{1}{N} \sum_{i=1}^k (x_i - \bar{x})^4}{\left(\frac{1}{N} \sum_{i=1}^k (x_i - \bar{x})^2 \right)^2} - 3$$

Covarianza

La covarianza es una medida de relación lineal entre dos variables aleatorias X_i y X_j , a mayor valor absoluto de la covarianza corresponde una mayor dependencia lineal entre X_i y X_j , valores positivos indican que cuando X_i crece también lo hace X_j , valores negativos indica que cuando X_i crece X_j decrece.

La covarianza de X_i y X_j se la estima de la siguiente forma:

$$\widehat{\text{COV}}(X_i, X_j) = \frac{\sum_{i=1}^p (X_i - \bar{x}_i)(X_j - \bar{x}_j)}{n-1}$$

$$i = 1, 2, \dots, p$$

$$j = 1, 2, \dots, p$$

donde \bar{x}_i y \bar{x}_j son los estimadores de los valores esperados de X_i y X_j respectivamente.

Coefficiente de correlación

Por la dificultad de utilizar la covarianza como una medida absoluta de la dependencia lineal porque su valor depende de la escala de medición y por consiguiente se dificulta determinar si una covarianza en particular es significativo o no. Se puede eliminar este problema estandarizando el valor de la covarianza, utilizando el coeficiente de correlación ρ_{ij} entre X_i y X_j , el cual se lo estima:

$$R_{ij} = \frac{\text{COV}(x_i, x_j)}{S_{x_i} S_{x_j}}$$

donde S_{x_i} y S_{x_j} son los estimadores de las desviaciones estándar de X_i y X_j respectivamente. Además el coeficiente de correlación entre dos variables aleatorias se encuentra entre -1 y +1; y entre más cercano esté el valor de ρ_{ij} hacia -1 o hacia 1 mayor será la relación lineal entre las variables.

Tabla de contingencia

La tabla de contingencia es un arreglo matricial de r filas y c columnas, donde r es el número de niveles del factor 1 o de la variables X_i y c el número de niveles del factor 2 o de la variable X_j , cada variable debe tener al menos dos niveles los cuales deben ser exhaustivas y mutuamente excluyentes. Las tablas de contingencia sirven para determinar la dependencia o independencia de dos variables aleatorias X_i y X_j . A continuación se muestra una tabla de contingencia:

		Factor 1			
Factor 2	Nivel 1	Nivel 2		Nivel c	
Nivel 1	X_{11}	X_{12}		X_{1c}	$X_{1.}$
Nivel 2	X_{21}	X_{22}		X_{2c}	$X_{2.}$
Nivel r	X_{r1}	X_{r2}		X_{rc}	$X_{r.}$
	$X_{.1}$	$X_{.2}$		$X_{.c}$	

Donde:

X_{ij} es el número de unidades de investigación sometidas al i -ésimo nivel del factor 2 y el j -ésimo nivel del factor 1.

$$X_{.i} = \sum_{j=1}^c X_{ij} \quad \text{y}$$

$$X_{.i} = \sum_{j=1}^r X_{ji}$$

El contraste de hipótesis planteada es:

H_0 : X_i y X_j son independientes

vs.

H_1 : $\neg H_0$

y éste se basa en:

$$E_{ij} = \frac{X_{.i} X_{.j}}{n} \quad \text{donde} \quad n = \prod_{i=1}^r \prod_{j=1}^c X_{ij}$$

$$X^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(X_{ij} - E_{ij})^2}{E_{ij}}$$

Se puede probar que X^2 bajo ciertas condiciones teóricas tiene una

distribución χ^2 con $(r-1)(c-1)$ grados de libertad, esto es:

$$X^2 > \chi^2(r-1)(c-1)$$

Bajo estas condiciones, se rechaza H_0 a favor de H_1 con $(1-\alpha)100\%$

de confianza si:

$$X^2 > \chi_{\alpha}^2(r-1)(c-1)$$

Análisis de Homogeneidad (HOMALS)

El análisis de homogeneidad, conocido como análisis de correspondencia múltiple, sirve para analizar una matriz de datos categóricos multivariada cuando se desea hacer un análisis de todas las variables a un nivel nominal.

HOMALS cuantifica los datos (categóricos) nominales mediante la asignación de valores numéricos a los casos (los objetos) y a las categorías. El análisis de homogeneidad se conoce también por el acrónimo HOMALS, del inglés homogeneity analysis by means of alternating least squares (análisis de homogeneidad mediante mínimos cuadrados alternantes).

El objetivo es describir las relaciones entre dos o más variables nominales en un espacio de pocas dimensiones que contiene las categorías de las variables así como los objetos pertenecientes a dichas categorías. Los objetos pertenecientes a la misma categoría se representan cerca los unos de los otros, mientras que los objetos de diferentes categorías se representan alejados los unos de los otros. Cada objeto se encuentra lo más cerca posible de los puntos de categoría para las categorías a las que pertenece dicho objeto.

El análisis de homogeneidad es similar al análisis de correspondencias, pero no está limitado a dos variables. Es por ello

que el análisis de homogeneidad se conoce también como el análisis de correspondencias múltiple. También se puede ver el análisis de homogeneidad como un análisis de componentes principales para datos nominales.

El análisis de homogeneidad es más adecuado que el análisis de componentes principales típico cuando puede que no se conserven las relaciones lineales entre las variables, o cuando las variables se miden a nivel nominal. Además, la interpretación del resultado es mucho más sencilla en HOMALS que en otras técnicas categóricas, como pueden ser las tablas de contingencia y los modelos loglineales. Debido a que las categorías de las variables son cuantificadas, se pueden aplicar sobre las cuantificaciones técnicas que requieren datos numéricos, en análisis subsiguientes.¹¹

“ Homals se aplica a tablas de contingencias en las que las que por filas se tienen n individuos y por columnas s variables categóricas con p_i $i=1, \dots, s$ categorías mutuamente excluyentes y exhaustivas.

La tabla de datos tiene; por lo tanto, la forma:

$$Z = [Z_1, Z_2, \dots, Z_s]$$

con Z , matriz $n \times p$, de forma que

$z_{ij} = 1$ si el individuo i -ésimo ha elegido la modalidad j .

$z_{ij} = 0$ si el individuo i -ésimo no ha elegido la modalidad j .

con $i=1, \dots, n$ y $j=1, \dots, p = p_1+p_2+\dots+p_s$

Esta técnica se basa en realizar un análisis de correspondencia sobre la llamada matriz de Burt:

$$\mathbf{B} = \mathbf{Z}'\mathbf{Z}$$

Las reglas de interpretación son:

- 1) Dos individuos están próximos si han elegido globalmente las mismas modalidades.
- 2) Dos modalidades están próximas han sido elegidas globalmente por el mismo conjunto de individuos.
- 3) La interpretación de los factores se hace teniendo en cuenta las contribuciones totales de cada variable que vienen dadas por

$$Ct_k(q) = \sum_{j \in q} Ct_k(j)$$

Se calculan también la varianza de las modalidades de cada variable y viene dadas por:

$$\sum_{j=1}^{p_q} n_{.j} b_{jk}^2 = \mu_k^2 Ct_k(q)$$

A ésta medida se la llama **medida de discriminación de la variable**.

- 4) $d^2 \left(\cdot, O \right) = \frac{n_{..}}{Z_{.j}} - 1$ por lo que una modalidad estará más alejada del

origen de coordenadas cuanto menor número de efectivos tenga.

- 5) La inercia de una variable $I(q) = \sum_{j=1}^{p_q} I(j) = \frac{1}{s} (p_q - 1)$ es función

creciente de su número de modalidades.

6) La inercia total vale $I = \frac{P}{s} - 1$ y no tiene ninguna significancia estadística.”¹⁰

Las herramientas utilizadas en el Análisis de Homogeneidad son:

Estadísticos y Gráficos: Los estadísticos que se obtienen del análisis de homogeneidad son: frecuencias, autovalores, historial de iteraciones, puntuaciones de objeto, cuantificaciones de categoría, medidas de discriminación. Entre las representaciones gráficas que brinda, están: gráficos de las puntuaciones de objetos, gráficos de las cuantificaciones de categoría y gráficos de las medidas de discriminación.

Consideración sobre los datos

“ *Datos*. Todas las variables son categóricas (nivel de escalamiento óptimo nominal). Utilice enteros para codificar las categorías. Para minimizar los resultados, utilice enteros consecutivos, comenzando por el 1, para codificar cada variable.

Supuestos. Todas las variables del análisis tienen cuantificaciones de categoría que pueden diferir para cada dimensión (nominal múltiple). En el análisis, sólo se utiliza un conjunto de variables. El número máximo de dimensiones utilizado en el procedimiento es el más pequeño entre el número total de categorías menos el número de variables sin datos perdidos y el número de casos menos 1. Por ejemplo, si una variable dispone de cinco categorías y la otra de

cuatro (sin datos perdidos), el número máximo de dimensiones es siete $((5+4) - 2)$. Si especifica un número superior al máximo, se utilizará el valor máximo.

Procedimientos relacionados. Para dos variables, el Análisis de homogeneidad es análogo al Análisis de correspondencias. Si piensa que las variables poseen propiedades ordinales o numéricas, se deben utilizar Componentes principales mediante escalamiento óptimo. Si hay conjuntos de variables que son de interés, se debe utilizar el Análisis de correlación canónica no lineal.”¹¹

Las propiedades básicas del Análisis de Homogeneidad son:

Las cuantificaciones de las categorías y las puntuaciones de objetos son representados en un espacio común.

Las soluciones sucesivas para las puntuaciones de los objetos no estén correlacionadas entre ellas, pero esto no implica que las cuantificaciones sucesivas de la misma variable sean no correlacionadas.

Existe una excepción a la regla anterior, si se aplica HOMALS a una situación con solamente dos variables categóricas, las cuantificaciones sucesivas de estas dos variables no serán correlacionadas para esto, existe otro programa denominado ANACOR.¹²

Vector aleatorio

Sean X_1, X_2, \dots, X_p p variables aleatorias sujetas a investigación. Se define un vector p variado $X \in R^p$, el que está compuesto por las p variables aleatorias como se muestra a continuación:

$$X^t = \begin{bmatrix} X_1 & X_2 & \dots & X_p \end{bmatrix}$$

Matriz de datos

En la matriz de datos X cada elemento x_{ij} representa el i -ésimo ente al cual se le realiza la j -ésima medida, cada columna corresponde a las p mediciones tomadas a un ente. Es decir, a n entes se les miden p características:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{p1} & x_{p2} & \dots & x_{pn} \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & \dots & X_n \end{bmatrix} \in R^p$$

X_1, X_2, \dots, X_n es una muestra tomada de una población de tamaño N que tiene p variables o características de interés (una población p variada).

Matriz de varianzas y covarianzas

Sea: $X^t = [X_1 \ X_2 \ \dots \ X_p]$ un vector p variado, se define para éste el estimador de la matriz de varianzas y covarianzas como:

$$\hat{\Sigma} = X \left(\mathbf{1}_n \mathbf{1}_n^t - \frac{1}{n} I_n \right) X^t$$

$$\hat{\Sigma} = \begin{bmatrix} \text{cov}(X_1, X_1) & \text{cov}(X_1, X_2) & \dots & \text{cov}(X_1, X_p) \\ \text{cov}(X_2, X_1) & \text{cov}(X_2, X_2) & \dots & \text{cov}(X_2, X_p) \\ \cdot & \cdot & \cdot & \cdot \\ \text{cov}(X_p, X_1) & \text{cov}(X_p, X_2) & \dots & \text{cov}(X_p, X_p) \end{bmatrix}$$

$$\hat{\Sigma} = \begin{bmatrix} S_{11}^1 & S_{12} & \dots & S_{1p} \\ S_{21} & S_{22}^1 & \dots & S_{2p} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ S_{p1} & S_{p2} & \dots & S_p^2 \end{bmatrix}$$

donde $S_{ij} = S_{ji}$, por lo tanto, $\hat{\Sigma}$ es asimétrica y por lo tanto diagonalizable ortogonalmente.

Componentes Principales

El análisis de componentes principales es una técnica estadística multivariada que permite la reducción de variables. En álgebra se lo considera como una combinación lineal de las p variables aleatorias observadas X_1, X_2, \dots, X_p , y para la geometría la combinación lineal representa la elección de un nuevo sistema de coordenadas, que fueron obtenidas al rotar el sistema original. Los ejes de este nuevo sistema representa la dirección de máxima variabilidad. Lo que nos

permite describir la estructura e interrelación de variables observables consideradas simultáneamente, determinando b combinaciones lineales de p variables originales que expliquen la mayor parte de la variación total, con lo que se consigue resumir y reducir las variables disponibles.

Sea $X^T = [X_1 \ X_2 \ \dots \ X_p]$ un vector aleatorio p-variado, y cada una de las variables que lo componen son variables aleatorias observables. El vector p-variado X tiene la matriz de varianzas y covarianzas Σ y sea $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ los valores propios correspondientes a Σ .

Consideremos las combinaciones lineales que se presentan a continuación:

$$Y_1 = a_1^T X = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \quad Y_2 = a_2^T X = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p$$

.

.

.

$$Y_b = a_b^T X = a_{b1}X_1 + a_{b2}X_2 + \dots + a_{bp}X_p$$

Entonces, Y_1, Y_2, \dots, Y_b son las componentes principales, cuyas varianzas tienen la siguiente característica $Var(Y_1) \geq Var(Y_2) \geq \dots \geq Var(Y_b) \geq 0$, y además las componentes no están correlacionadas.

Se puede demostrar que:

$$\text{Var}(Y_i) = a_i^T \Sigma a_i, \text{ para } i = 1, 2, \dots, b$$

$$\text{Cov}(Y_i, Y_j) = a_i^T \Sigma a_j, \text{ para } j = 1, 2, \dots, b$$

y deben cumplir con $\|a_i\| = 1$ para $i = 1, 2, \dots, p$ y $\langle a_i, a_j \rangle = 0$ para $i \neq j$ (ai y aj son ortonormales).

Donde

$\|a_i\|$ es la norma del vector ai

y

$\langle a_i, a_j \rangle$ es el producto interno

entre los vectores ai y aj.

➤ La primera componente principal es la combinación lineal $Y_1 = a_1^T X$

de máxima varianza, donde $\|a_1\|$ es unitaria.

➤ Para la i-ésima componente principal es la combinación lineal

$Y_i = a_i^T X$, que maximiza la varianza de Y_i , que tiene norma unitaria

para i y las $\text{Cov}(Y_i, Y_k) = 0$ para $k < i$.

➤ Los resultados obtenidos son:

Σ es la matriz de varianzas y covarianzas asociadas con el vector

aleatorio $X^T = [X_1 \ X_2 \ \dots \ X_p]$,

Σ tiene los pares de valores propios y sus correspondientes vectores

propios $(e_1, e_1), (e_2, e_2), \dots, (e_p, e_p)$ donde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$.

Entonces, se puede probar que la i -ésima componente principal viene dada por:

$$Y_i = e_i^T X = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p \text{ para } i = 1, 2, \dots, b$$

$$\text{La } \text{Var}(Y_i) = e_i^T \Sigma e_i = \lambda_i, \text{ para } i = 1, 2, \dots, b$$

$$\text{Cov}(Y_i, Y_k) = e_i^T \Sigma e_k \text{ para } i \neq k$$

El porcentaje total de la varianza contenida por la i -ésima componente principal, o su explicación está dado por:

$$\frac{\lambda_i}{\sum_{i=1}^b \lambda_i}$$

Muestreo Estratificado

Se estratifica cuando es posible dividir una población heterogénea en subpoblaciones, en las que cada una sea internamente homogénea.

En este tipo de muestreo, la población de N unidades, se particiona primero en subpoblaciones de N_1, N_2, \dots, N_k unidades, respectivamente.

Estas subpoblaciones, en conjunto comprenden a toda la población. por lo tanto:

$$N_1 + N_2 + \dots + N_k = N$$

Las subpoblaciones se denominan estratos, para obtener todo el beneficio de la estratificación, el tamaño de los estratos debe ser conocido. Una vez determinados los estratos, se extrae una muestra de cada uno, las extracciones deben hacerse independientemente.

Los tamaños de muestras dentro de los estratos se denominan con n_1, n_2, \dots, n_k respectivamente, tales que $n =$ tamaño de la muestra =

$$\sum_{i=1}^k n_i .$$

Si la estratificación se obtiene seleccionando una muestra aleatoria simple en cada estrato de forma independiente, el muestreo se denomina muestreo aleatorio estratificado, pero en general nada impide utilizar diferentes tipos de selección en cada estrato.

Las principales razones para utilizar el método de muestreo estratificado son:

Dar estimaciones separadas para ciertas subpoblaciones, otra de las razones es agrupar unidades de muestreo homogéneas entre sí en estratos, con el objeto de mejorar la precisión de las estimaciones globales, por ejemplo: en poblaciones humanas, las personas que viven en instituciones (como hoteles, hospitales, cárceles) se colocan en un estrato diferente de las que viven en casa ordinarias. Si se toma una muestra aleatoria simple de toda la población sin considerar los estratos (o sus diferencias), puede ser muy probable que se incluya un número demasiado escaso de algunos de los estratos y muy excesivos de otros.

Estratificar da lugar entonces, a una ganancia en precisión de las estimaciones de características de la población total.

Afijación de la muestra.- Es la asignación del tamaño muestral n entre los diferentes estratos, es decir, se determinan los n_h (tamaños en cada estrato) que verifiquen:

$$\sum_{i=1}^k n_h = n_1 + n_2 + \dots + n_k = n$$

Los tipos de afijaciones son: la **afijación uniforme** en la que se toman todos los n_h iguales. Este tipo de afijación implica dar la misma importancia a todos los estratos, en cuanto a tamaño y favoreciendo a los pequeños en cuanto a precisión. La **afijación proporcional** en la que se asigna a cada estrato, un número de elementos en la muestra proporcional al tamaño de cada estrato.

$$n_h = \frac{n}{N} N_h$$

donde N_h es el tamaño de cada estrato.

Si se hace una afijación proporcional, $\bar{x}_{est} = \bar{X}$. La varianza del estimador de la media sería:

$$Var(\bar{x}_{est}) = \sum_{h=1}^L \frac{n_h^2}{n^2} \left(1 - f \right) \frac{S_h^2}{n_h}$$

en la que $f = n/N$

Afijación de mínima varianza consiste en determinar valores de n_h de manera que para un tamaño de muestra fijo igual a n la varianza sea mínima.

El problema reside en hacer mínima la expresión $V(\bar{x}_{est})$ bajo la

condición: $\sum_h^L n_h = n$

Así se observa, que los valores de n_h son proporcionales a los productos $N_h S_h$.

$$n_h = n \frac{N_h S_h}{\sum_{h=1}^L N_h S_h}$$

Si bien, la afijación proporciona la mínima varianza del estimador \bar{x}_{est} , también requiere mejor conocimiento de las características poblacionales, ya que no basta saber el tamaño de la muestra de cada estrato sino conocer también su cuasivarianza S_h , pero debido a que ésta es regularmente desconocida, se la estima a partir de una muestra piloto.

La utilidad de la afijación de mínima varianza es mayor si hay grandes diferencias de variabilidad de los estratos. Por último, la **afijación óptima** es la que minimiza la varianza del estimador considerando los costos de cada estrato.

En la presente investigación se utilizará afijación proporcional.

2.3 Definiciones Básicas Financieras y Económicas

Activos circulantes

Activos a corto plazo, que se espera que se conviertan en efectivo en un plazo de un año o menos.

Balance General

Declaración resumida de la situación financiera de la empresa en un momento dado.

Capital de trabajo

Activos circulantes que representan la proporción de inversión que circula de una forma a otra en el conducto ordinario de negocios.

Depreciación

Cargo sistemático de una parte de los costos de activos fijos contra los ingresos anuales a través del tiempo.

Estado de resultados

Proporciona un resumen financiero de los resultados de las operaciones de la empresa durante un período determinado

Gasto de capital

Desembolso de fondos hecho por la empresa que se espera que produzca beneficios durante un período de más de un año.

Gasto operativo

Fondos que desembolsa la empresa que dan como resultado beneficios recibidos dentro de un año

Ingresos Netos

Fondos recibidos realmente de la venta de un valor

Inversión inicial

El flujo negativo de efectivo relevante para un proyecto propuesto en el tiempo cero.

Liquidez

Capacidad de una empresa para satisfacer obligaciones a corto plazo conforme se venzan.

Pasivo circulante

Pasivos a corto plazo, que se espera que se paguen en un plazo de un año o menos.

Tasa interna de rendimiento

Técnica sofisticada para preparar presupuestos de capital; la tasa anual compuesta de rendimiento es lo que la empresa ganará si invierte en el proyecto y recibe los flujos positivos de efectivo dados.

Valor Presente Neto

Se obtiene sustrayendo la inversión inicial de un proyecto del valor presente de sus flujos positivos de efectivo descontados a una tasa equivalente al costo del capital de la empresa.

2.4 Descripción y Codificación de Variables a ser utilizadas

En el presente estudio se tiene un total de 48 variables de carácter cuantitativo y cualitativo. A continuación se hace una descripción de cada una de ellas y se presenta la respectiva codificación.

2.4.1 Descripción de variables de “Información General”

Variable IG₁= Género

La variable cualitativa binomial de carácter nominal IG₁ representa el género de las personas entrevistadas. Esta variable puede tomar dos valores posibles: Masculino y femenino.

Variable IG₂= Nivel de Mando

Se identifica a esta variable como el nivel de mando de los empresarios entrevistados. El nivel de mando de estas personas puede ser: Vicepresidente, Gerente General, Gerente de área, Jefes de departamentos u Otros.

Variable IG₃= Relación que tiene con la empresa

Esta variable indica la relación que el empresario tiene con la empresa, esto puede ser: Dueño, Accionista o No Accionista.

Variable IG₄ = Tiempo trabajando en la Empresa

Esta variable permitirá indagar cuanto tiempo tiene trabajando el empresario entrevistado, esta variable puede tomar los siguientes rangos de valores: Menos de 1 año, de 1 a 3 años, de 3 a 6 años y Más de 6 años.

Variable IG₅ = Número de empleados de la empresa

Esta variable cuantitativa representa el número de empleados de la empresa.

Variable IG₆ = Año en que empezó a operar la empresa

A través de esta variable se pretende conocer el año en el que empezó a funcionar la empresa.

Variable IG₇ = Sucursal de empresa

Con esta variable se pretende conocer si la empresa tiene sucursales en Guayaquil o en otra ciudad.

Variable IG₈ = Ciudad de Sucursal

Esta variable cualitativa tiene como objetivo determinar la ciudad en las que las empresas entrevistadas tienen sucursal.

2.4.2 Descripción de variables de “Sección Exploratoria”

Variable SE₉ = Instalar impresora

Esta variable pretende conocer la opinión de los empresarios acerca de la siguiente proposición 1: “Es apropiado instalar una impresora de oficina a un Terminal de Información”.

Variable SE₁₀ = Acceso a Internet

La variable SE₉ representa la opinión acerca de las proposición 2: “Considera Ud. que las Terminales deben tener acceso a Internet.”

Variable SE₁₁ = Personas autorizadas

Es un variable que nos indica si es conveniente autorizar a personas para encender o apagar terminales a través de la proposición 3: “Es conveniente que las Terminales de Información solo deben ser encendidas o apagadas por personas autorizadas.”

Variable SE₁₂ = Agilidad en transacción

Esta variable pretende analizar la grado de aceptación con la proposición 4: “Los Terminales de Información brindan mayor agilidad al usuario al momento de realizar una transacción.”

Variable SE₁₃ = Desarrollo tecnológico

A través de esta variable se desea conocer el criterio de los empresarios respecto al desarrollo tecnológico del país, esto es

indicado en la proposición 5: “La implementación de Terminales de Información ayuda al desarrollo tecnológico del país.”

Variable SE₁₄ = Nivel de educación

Con esta variable se intenta indagar cual es la calificación que los empresarios le asignan a la proposición 6: “El nivel de educación de los ecuatorianos influye en el uso adecuado de Terminales”

Variable SE₁₅ = Medio de publicidad

Esta variable pretende analizar la grado de aceptación con la proposición 7: “Los Terminales de Información son un medio de publicidad para las empresas”

Variable SE₁₆ = Recurso humano competente

La variable SE₁₅ representa la opinión acerca de las proposición 8: “El Recurso Humano ecuatoriano es competente para proporcionar la instalación y mantenimiento de terminales de información.”

Variable SE₁₇ = Desarrollo de turismo

La variable SE₁₆ representa el acuerdo de las personas entrevistadas con la proposición 9: “Los Terminales de Información ayudan al desarrollo del Turismo en el Ecuador.”

Variable SE₁₈ = TI instalados en la calle

La Variable SE₁₇ se refiere a la proposición 10: “Los Terminales de Información pueden ser instalados en la calle”

Variable SE₁₉ = Guías en grandes establecimientos

Esta variable pretende analizar la grado de aceptación con la proposición 11:” Los Terminales de Información son utilizados como guías para el cliente dentro de grandes establecimientos comerciales.”

2.4.3 Descripción de las variables de “Características de la investigación”**Variable CI₂₀ = Implementar TI en la organización.**

Esta variable indica si las empresas están dispuestas a implementar Terminales de información en su organización, se lo expresa en la proposición 12: “Esta organización necesita implementar terminales de información”

Variable CI₂₁ = Objetivo servicio al cliente

Esta variable pretende captar cual es la aceptación a la proposición 13: “Los terminales de Información se implementan con el objetivo de mejorar el servicio al cliente de la organización.”

Variable CI₂₂ = Objetivo avance tecnológico

La variable SE₂₁ representa el acuerdo de las personas entrevistadas con la proposición 14: “Los terminales de Información se implementan dentro de la organización con el objetivo de ser parte del avance tecnológico del país.”

Variable CI₂₃ = Reducción costo personal

La variable SE₂₂ representa la opinión acerca de las proposición 15: “La instalación de Terminales de Información reduce el costo de personal”.

Variable CI₂₄ = Reducirá puestos de trabajo

Esta variable indica la opinión de los empresarios respecto a la proposición 16: “La implementación de Terminales reducirá los puestos de trabajo dentro de su organización”.

Variable CI₂₅ = Atención a través de tickets

Esta variable pretende analizar la grado de aceptación con la proposición 17: “Los Terminales de Información se utiliza para atender a los clientes en el respectivo orden de llegada a través de ticket”.

Variable CI₂₆ = Saldo de cuenta

La variable SE₂₅ representa la opinión acerca de la proposición 18: “Los Terminales de Información sirven al cliente para conocer el saldo de su cuenta.”

Variable CI₂₇ = Atención desde lugar externo

Esta variable pretende analizar la grado de aceptación con la proposición 19: “Los Terminales de Información sirven al cliente para obtener un turno para la atención de algún departamento desde un lugar distinto a la empresa.”

Variable CI₂₈ = Información de descuentos y promociones

Esta variable representa la calificación asignada a la proporción 20: “Los Terminales de Información sirven al cliente para informarse de los descuentos y promociones de los productos de la empresa.”

Variable CI₂₉ = Implementar TI en servicio al cliente

A través de esta variable se desea conocer el criterio de los empresarios respecto a implementa terminales de información en el área de servicio al cliente.

Variable CI₃₀ = Implementar TI en Recurso Humano

Esta variable binomial intenta descubrir si las organizaciones necesitan implementar terminales de información en el área de Recurso Humano.

2.4.4 Descripción de variables de “Información Financiera”

Variable IF₃₁ = Inversión en tecnología

Esta variable pretende indagar cual es la inversión que las empresas dedican a la tecnología anualmente.

Variable IF₃₂ = Comprar TI

La Variable SE₃₁ se relaciona a la pregunta: Considera Ud. que la empresa estaría dispuesta a comprar Terminales de Información?

Variable IF₃₃ = Cuántas TI

Esta variable representa la cantidad de terminales de información que las empresas desearían comprar.

Variable IF₃₄ = Precio de compra

Esta variable representa el precio en rangos que las empresas estarían dispuestas a pagar por la compra de un Terminal de información.

Variable IF₃₅ = Costo de mantenimiento

En caso de que las empresas compren terminales de información es necesario conocer cual es el precio que estarían dispuestos pagar por el mantenimiento de estas.

Variable IF₃₆ = Alquiler TI

La variable IF₃₅ indica si las empresas estarían dispuestas a alquilar terminales de información.

Variable IF₃₇ = Cantidad de TI para alquilar

Esta variable se refiere a la cantidad de terminales de información que las empresas desean alquilar.

Variable IF₃₈ = Precio de alquiler

Esta variable representa el precio en rangos que las empresas estarían dispuestas a pagar por el alquiler semanal de un Terminal de información.

Variable IF₃₉ = Tiempo de alquiler

Esta variable indica el tiempo de alquiler de los terminales de información.

Variable IF₄₀ = Frecuencia de alquiler

Esta variable se refiere a las veces que las empresas desean alquilar los terminales de información en el año.

2.4.5 Descripción de variables de “Información Adicional”**Variable IA₄₁ = Conoce a empresas de TI**

La variable IA₄₀ pretende indagar sobre el conocimiento de los empresarios de empresas que se dediquen a terminales de información.

Variable IA₄₂ = Empresa nueva

Esta variable intenta conocer si las empresas estarían dispuestas a confiar servicios de terminales a una empresa nueva.

Variable IA₄₃ = Cumplimiento

Mediante esta variable se pretende conocer el orden de importancia que le asignan a la característica cumplimiento.

Variable IA₄₄ = Precios

A través de esta variable se trata de percatarse el orden de importancia que le asignan a la característica precio.

Variable IA₄₅ = Confiabilidad

Esta variable indica el orden de importancia de la característica Confiabilidad.

Variable IA₄₆ = Garantía

Mediante esta variable se pretende conocer el orden de importancia que le asigna a la característica Garantía.

Variable IA₄₇ = Servios Post-ventas

Es una variable que nos da a conocer el orden de importancia de la característica Servios Post-ventas.

Variable IA₄₈ = Diseño

Esta variable pretende indagar cual es el orden de importancia de la característica diseño.

Variable IA₄₉ = Patrocinador

En esta variable se pretende conocer si las empresas estarían dispuestas a participar como patrocinador o utilizar nuestros servicios para participar como auspiciantes, es decir que utilicen los terminales de información como medio de publicidad, colocando sus logos en estas.

2.5 Técnicas de Muestreo y Selección de la Muestra**2.5.1 Determinación de la población objetivo y Marco Muestral**

La población objetivo de esta investigación son las 247 empresas que cumplen con el siguiente perfil:

- Son empresas de servicio
- Manejan muchos datos
- Tienen alto rendimiento o liquidez

En estas empresas, implementar Terminales de Información facilitará el trabajo de transacciones relacionadas al consumidor o al recurso humano. Además el alto rendimiento o liquidez de estas empresas garantizar la validez de la información referente a

la compra de terminales de información puesto que estos tienen precios altos que serían menos accesibles a empresas que no cumplan con esta última característica.

El marco muestral utilizado está conformado por el listado de empresas de Guayaquil catalogadas como las más importantes proporcionado por la Superintendencia de Compañías, cuya categorización está realizada en base a variables como rentabilidad, liquidez, ingresos por ventas, patrimonio y activos.

Esta información se encuentra segmentada por tipo de actividad económica lo que se muestra en la tabla # 1.

Las Unidades de Investigación serán las empresas que cumplan con el perfil referido anteriormente.

Tabla 1 Tipo de actividad económica de empresas de Guayaquil	
<i>Tipo de actividad económica</i>	<i>Número de empresas</i>
Comercio, restaurantes y hoteles	155
Transportes, almacenamiento y Comunicaciones	22
Establecimientos financieros, seguros, bienes inmuebles, y servicios prestados a las empresas	59
Servicios comunales, sociales y personales	11
Total	247

Fuente: Superintendencia de Compañías

Para efecto de este estudio se escogieron las empresas que corresponden a las siguientes actividades económicas: Comercio, Transporte y Comunicaciones, Servicios personales y Servicios empresariales puesto que estas son conforman el sector económico servicios definidos según la segunda revisión de la Clasificación Internacional Industrial Uniforme (CIIU) de las Naciones Unidas.

2.5.2 Diseño Muestral

El diseño muestral de esta investigación se basa en la utilización del muestreo aleatorio estratificado, se segmentó la población objetivo en cuatro estratos que corresponden a las actividades económicas: Comercio, Transporte y Comunicaciones, Servicios personales y Servicios empresariales.

Se utilizará el criterio de afijación proporcional para determinar cuantas unidades de la muestra se distribuyen proporcionalmente en cada estrato, esto es el cociente entre el tamaño de cada estrato y la población objetivo

2.5.3 Diseño de cuestionario

El instrumento de medición es el cuestionario, con el cual se alcanzarán los objetivos determinados para la investigación.

En el cuestionario consta de cinco secciones relacionadas a la investigación:

1. Información general
2. Sección exploratoria
3. Características de la investigación
4. Información financiera
5. Información adicional

La primera sección esta diseñada para obtener información de la persona entrevistada: género, nivel de mando y relación con la empresa, además datos generales de la empresa como el año de inicio, número de empleados y sucursales, la segunda sección esta planteada como una sección exploratoria, para conocer la percepción de los empresarios sobre los terminales de información, la tercera sección busca indagar sobre características específicas de la investigación, sobretodo la necesidad de terminales de información que existe en la empresa; la información financiera de la empresa y valores relevantes para el proyecto es lo que se trata de obtener en la cuarta sección y en la última sección se pretende conocer aspectos importantes respecto a SINTEC, que servirán como apoyo para su creación.

2.5.4 Determinación de la muestra

2.5.4.1 Tamaño de la muestra

Uno de los primeros pasos para obtener el tamaño de la muestra real de la investigación es la extracción de una muestra piloto. Se tiene que la población objetivo son las 247 empresas que poseen las características específicas para la investigación, de esta se seleccionó a 24

empresarios, bajo el supuesto de muestreo aleatorio simple entre los diversos tipos de actividades económicas. La variable a analizar es la variable X_5 (esta organización necesita implementar terminales de información) que es nuestra variable de interés, la que se expresó en forma de proposición con una escala de 1 a 5, se escogió la zona de acuerdo (calificación 4 o 5) para obtener la proporción p ; obteniendo como resultado 0.83. Tomando un error del 0.06 y con una confianza del 95% se procede a calcular el tamaño de la muestra como se detalla a continuación:

$$N_0 = \frac{Z_{\alpha/2}^2 * p * q}{E^2} = \frac{1.96^2 * 0.83 * 0.17}{0.06^2} = 148.21$$

$$n = \frac{N_0 * N}{N_0 + N} = \frac{148.21 * 247}{148.21 + 247} = 92.63 \cong 93$$

donde:

N = tamaño de la población conocida

$Z_{\alpha/2}^2$ = coeficiente de confianza

p = proporción de acuerdo con la proposición

$q = 1 - p$

E^2 = error de estimación

2.5.4.2 Selección de la muestra

Para determinar cuántos empresarios entrevistar exactamente debemos seleccionar por tipo de actividad

económica, entonces se procede a aplicar afijación proporcional, tal como se muestra en la tabla 2.

Tabla 2
Proporción de empresas a entrevistar por tipo de actividad económica

Estratos (actividades económicas)	Numero de empresas	Proporción W_{ij}	Empresas a entrevistar
<i>Comercio</i>	155	0,63	58
<i>Transportes y Comunicaciones</i>	22	0,09	9
<i>Servicios a Empresas</i>	59	0,24	22
<i>Servicios Personales</i>	11	0,04	4
<i>Total</i>	247	1.00	93