

**ESCUELA SUPERIOR POLITÉCNICA DEL
LITORAL**

Instituto de Ciencias Matemáticas

**“ANÁLISIS ESTADÍSTICO DEL CULTIVO Y PRODUCCIÓN DE
LA CAÑA DE AZÚCAR”**

TESIS DE GRADO

Previa la obtención del Título:

**INGENIERO EN ESTADÍSTICA -
INFORMÁTICA**

Presentada por:

JESSICA MARIELA CARVAJAL MORALES

Guayaquil - Ecuador

2003

AGRADECIMIENTO

Al Mat. John Ramírez Figueroa, Director de la Tesis por su invaluable ayuda.

Al Ingenio Valdez de Milagro, por las facilidades prestadas en cuanto a información, lo que me ha permitido culminar con gran éxito el desarrollo de esta tesis.

DEDICATORIA

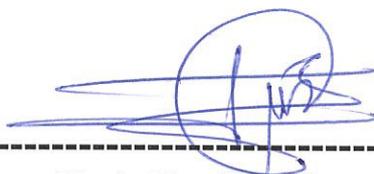
A Dios: por ser siempre bondadoso con
migo

A mis padres Ena y Gualberto: Por darme
siempre su apoyo y confianza y todo su
amor.

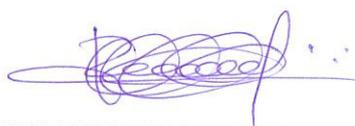
A mis hermanas: Por estar siempre a mi
lado

A Edwin: Por ser comprensivo y estar
siempre junto a mi en mis éxitos y fracasos.

TRIBUNAL DE GRADUACION



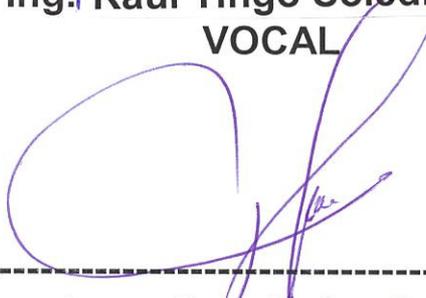
Mat. Medina Sancho
Director del Instituto de Ciencias Matemáticas



Mat. John Ramírez Figueroa
DIRECTOR DE TESIS



Ing. Raúl Tingó Soledispa
VOCAL



Ing. Jorge Fernández Ronquillo
VOCAL

DECLARACION EXPRESA

“La responsabilidad del contenido de esta Tesis de Grado, me corresponden exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITECNICA DEL LITORAL”.

(Reglamento de Graduación de la ESPOL).

A handwritten signature in black ink, appearing to read 'Jéssica Carvajal', written over a horizontal line.

JÉSSICA MARIELA CARVAJAL MORALES

RESUMEN

La presente tesis es el análisis estadístico de la producción y cultivo de caña de azúcar. Nuestro país es conocido como productor de azúcar, por cual intervienen variables agrícolas en la producción y cultivo de la caña.

En el primer capítulo se expone como introducción un resumen de la fase agrícola y Producción de la caña de azúcar.

En el segundo Capitulo se revisa el marco teórico que será utilizado, comprende el análisis univariado y el análisis por componentes principales.

En el tercer capitulo se detallan las variables que serán estudiadas, y se muestra su estadística descriptiva. Por medio del análisis multivariado se observan las diferentes relaciones entre las variables y agrupaciones que intervienen en la producción y cultivo de la caña de azúcar.

En el cuarto y ultimo capitulo presentaremos las conclusiones y recomendaciones que se han obtenido.

ABREVIATURAS Y SIMBOLOGIAS

ABREVIATURAS

AC= Antes de Cristo

TM,tons= Tonelada Métricas

Has= Hectáreas

RENDIT= Rendimiento

TMC/HAS= Toneladas Métricas por Hectáreas en Rendimiento

TM-CAÑA= Toneladas Métricas en producción

Kgrs,Kg = Kilogramos

US\$= Dólares

Cm=Centímetro

°c=Grados

X₁ =Superficie Sembrada

X₂ =Superficie Cosechada

X₃ =Producción

X₄ =Roza de Cantero

X₅ =Herbicida

X₆ =Rastrillo

X₇ =Escarificado

X₈ =Riego

X₉ =Sacada de Paja cauca

X₁₀ =Aporque

X₁₁ =Fertilización

SIMBOLOGIAS

\bar{x} = Media Aritmética

x_1, x_2, \dots, x_n = Elementos de una muestra aleatoria de tamaño n

$\text{máx}(x_1, x_2, \dots, x_n)$ = Máximo valor de los n elementos

s^2 = Varianza

$\hat{F}(X)$ = Función Empírica

$x_{(k)}$ = Estadístico de Orden

$X(ij)$ = Observación de la j-ésima variable en el i-ésimo individuo

p_1, p_2, \dots, p_{99} = Percentiles

D_1, D_2, \dots, D_9 = Deciles

Q_1, Q_2, Q_3 = Quartiles

H_0, H_1 = Prueba de Hipótesis

$E(X_j)$ = Vector de medias

$S(j) = 1/n \sum_1^n (X_{ij} - \bar{X}(j))^2$ = Varianza muestral de la j-ésima variable

$S_{ik} = 1/n \sum_1^n (X_{ij} - \bar{X}(j))(X_{ik} - \bar{X}(k))$ = Covarianza entre la j-ésima y la k-ésima variable

$S = E[(X - \mu)(X - \mu)^T]$, $\text{Cov}(x)$ = Matriz de varianza y covarianza

$r(jk) = \frac{S(jk)}{\sqrt{S(j)S(k)}} = \frac{S(jk)}{S(j)S(k)}$ = Coeficientes de correlación entre la j-ésima y

la k-ésima variable

$r(jk)$ = Correlación de la j-ésima fila con la k-ésima columna

$X = X_1, X_2, \dots, X_p$ = vector aleatorio

$X^T = (X_1, X_2, \dots, X_p)$ = Traspuesta de un vector x

$\lambda_1 \geq \lambda_2 \geq \dots \lambda_p \geq 0$ = Vectores propios

$\det(\Sigma - \lambda I)$ = Determinante de la resta de dicha matriz

Y_1, Y_2, \dots, Y_p = P variables

$\sigma_{ii} = \sigma_i^2 = \text{Var}(X_i)$, $\text{Var}(Y_i) = \text{Var}(d_i^T X) = d_i^T \Sigma d_i$ = Varianza

$\text{Cov}(Y_i, Y_k) = d_i^T \Sigma d_k$ = Covarianza

$d_i^T X$, $Y_i = e_i^T X = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p$ = La i-ésima componente

ρ , R = Matriz de correlación

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

$\max_a / e_1, e_2, \dots, e_k [a^T \Sigma a / a^T a] = \lambda_{k+1}$ = Máximo de los valores característicos

$(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$ = Valores característicos

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ = Valores Característicos

Σ = Sumatoria, MATRIZ

e_i = E-igenvectores

R^n = Dimensión

μ = Vector de Medias

INDICE GENERAL

	Pág.
AGRADECIMIENTO.....	II
DEDICATORIA.....	III
DECLARACIÓN EXPRESA.....	V
RESUMEN	VI
INDICE GENERAL	VII
INDICE CONTENIDO.....	VII
INDICE DE TABLAS	XIII
INDICE DE ILUSTRACIONES	XVII

INDICE CONTENIDO

INTRODUCCIÓN.....	22
CAPITULO I.....	23
1. CULTIVO Y PRODUCCIÓN DE LA CAÑA DE AZÚCAR EN EL ECUADOR.....	23
1.1 Generalidades de la caña de azúcar.....	24
1.2 Antecedentes Históricos en el Ecuador.....	25
1.3 FASE AGRICOLA.....	29

1.3.1 PROCESOS QUE SE NECESITAN PARA REALIZAR LA SIEMBRA DE CAÑA DE AZUCAR.....	29
1.3.1.2 SEMILLEROS DE CAÑA.....	29
1.3.1.3 SELECCIÓN DE LA SEMILLA.....	32
1.3.2 SIEMBRA	33
1.3.3 ETAPAS DE LA CAÑA DE AZUCAR DURANTE SU RECIMIENTO	34
1.3.4 COSECHA.....	34
1.3.5 PRINCIPALES ENFERMEDADES, PLAGAS Y MALEZAS DE LA CAÑA DE AZUCAR.....	40
1.3.5.1 ENFERMEDADES	40
1.3.5.2 PLAGAS	40
1.3.5.3 MALEZAS	42
2. CULTIVO DE LA CAÑA DE AZUCAR EN EL ECUADOR.....	43
2.1 SUPERFICIE SEMBRADA	43
2.2 CAPACIDAD INSTALADA POR INGENIO	45
2.2.1 PRECIOS A NIVEL DE FINCA.....	48
2.3 PRODUCCION DE LA CAÑA DE AZÚCAR.....	51
CAPITULO II.....	54

DESCRIPCION DE LAS TECNICAS ESTADISTICAS A UTILIZAR.....	54
INTRODUCCIÓN.....	54
2.1 Análisis univariado.....	54
2.1.1 HISTOGRAMA DE FRECUENCIAS.....	57
2.1.2 OJIVA.....	58
2.1.3 DIAGRAMA DE CAJAS.....	60
2.1.4 PRUEBA DE BONDAD DE AJUSTE: KOLMOGOROV-SMIRNOV.....	61
2.2 ANALISIS MULTIVARIADO.....	62
2.2.1 DEFINICIONES.....	63
2.2.2 MATRIZ DE DATOS.....	63
2.2.3 VECTOR DE MEDIAS Y MATRICES DE VARIANZA Y COVARIANZAS.....	64
2.2.4 MATRIZ DE CORRELACIÓN.....	67
2.2.5 EL METODO DE ANALISIS POR COMPONENTES PRINCIPALES.....	68
CAPITULO III.....	76
ANÁLISIS ESTADÍSTICO.....	76
INTRODUCCIÓN.....	76

3.1 Variables a utilizar	76
3.1.1 Superficie Sembrada	76
3.1.2 Superficie Cosechada	77
3.1.3 Producción	77
3.1.4 Roza de Canteros.....	77
3.1.5 Herbicidas	78
3.1.6. Rastrillo	78
3.1.7 Escarificado	78
3.1. 8 Riego.....	79
3.1.9 Sacada de paja cauca.....	79
3.1.10 Aporque	80
3.1.11 Fertilización.....	80
3.2 ESTADÍSTICA DESCRIPTIVA DE LAS VARIABLES DE ESTUDIO	80
3.2.1 Superficie Sembrada	80
3.2.2 SUPERFICIE COSECHADA	85
3.2.3 PRODUCCIÓN	89
3.2.4 ROZA DE CANTEROS.....	92
3.2.5 INDICADOR DE ROZA DE CANTERO.....	95
3.2.6 HERBICIDAS	98
3.2.7 INDICADOR DE HERBICIDA.....	100
3.2.8 RASTRILLO	103

3.2.9 INDICADOR DE RASTRILLO	105
3.2.10 ESCLARIFICADO	108
3.2.11 INDICADOR DE ESCLARIFICADO	110
3.2.12 RIEGO	113
3.2.13 INDICADOR DE RIEGO	115
3.2.14 SACADA DE PAJA CAUCA	117
3.2.15 INDICADOR DE SACADA DE PAJA CAUCA	120
3.2.16 APORQUE.....	123
3.2.17 INDICADOR DE APORQUE	125
3.2.18 Fertilización.....	128
3.2.19 INDICADOR DE FERLIZACIÓN.....	130
3.3 ANALISIS MULTIVARIADO DE LAS VARIABLES DE INVESTIGACIÓN	133
3.3.1 Análisis de Componentes Principales	133
CAPITULO IV	150
CONCLUSIONES Y RECOMENDACIONES	150
4.1 Conclusiones	150
4.2 Recomendaciones.....	153
ANEXOS	155
BIBLIOGRAFIA	159

INDICE DE TABLAS

TABLA I: SEMILLEROS POR HECTAREA.....	31
TABLA II: ENFERMEDADES DE LA CAÑA DE AZUCAR	40
TABLA III: PRODUCCIÓN DE AZUCAR DEL ECUADOR.....	43
TABLA IV: HECTAREA SEMBRADA DE LOS INGENIOS DEL ECUADOR DE Junio/2000 -Junio /2001	44
TABLA V: CAPACIDAD INSTALADA.....	47
TABLA VI: PRECIOS NOMINALES DE LA CAÑA DE AZUCAR DE ENERO A OCTUBRE DE 1998	49
TABLA VII: PRECIOS PARA LOS CAÑICULTORES Años 1999-2000 (US\$/TM)	51
TABLA VIII: Estadística descriptiva: Variable Superficie Sembrada	81
TABLA IX : TABLA DE FRECUENCIAS DE SUPERFICIE SEMBRADA	82
TABLA X: Estadística descriptiva: Variable Superficie cosechada	86
TABLA XI TABLA DE FRECUENCIAS DE SUPERFICIE COSECHADA	86
TABLA XII: Estadística descriptiva: Variable Producción.....	89
TABLA XIII: TABLA DE FRECUENCIAS DE LA PRODUCCIÓN.....	90
TABLA XIV: ESTADÍSTICA DESCRIPTIVA: VARIABLE ROZA DE CANTERO.....	92
TABLA XV: TABLA DE FRECUENCIAS DE ROZA DE CANTERO.....	93
TABLA XVI: ESTADÍSTICA DESCRIPTIVA: VARIABLE ROZA DE CANTERO MEDIANTE UN INDICADOR DE VECES.....	95
TABLA XVII: TABLA DE FRECUENCIA DE ROZA DE CANTERO MEDIANTE UN INDICADOR DE VECES	96

TABLA XVIII: ESTADÍSTICA DESCRIPTIVA: VARIABLE HERBICIDAS	98
TABLA XIX: TABLA DE FRECUENCIA DE HERBICIDAS.....	99
TABLA XX: ESTADÍSTICA DESCRIPTIVA: VARIABLE HERBICIDAS MEDIANTE UN INDICADOR DE VECES	100
TABLA XXI: TABLA DE FRECUENCIA DE HERBICIDAS MEDIANTE UN INDICADOR DE VECES	101
TABLA XXII: ESTADÍSTICA DESCRIPTIVA: VARIABLE RASTRILLO	103
TABLA XXIII: TABLA DE FRECUENCIAS DE RASTRILLO	104
TABLA XXIV: ESTADÍSTICA DESCRIPTIVA: VARIABLE RASTRILLO MEDIANTE UN INDICADOR DE VECES	105
TABLA XXV: TABLA DE FRECUENCIA DE RASTRILLO MEDIANTE UN INDICADOR DE VECES	106
TABLA XXVI: ESTADÍSTICA DESCRIPTIVA: VARIABLE ESCARIFICADO.....	108
TABLA XXVII: TABLA DE FRECUENCIA DE ESCARIFICADO	109
TABLA XXVIII: ESTADÍSTICA DESCRIPTIVA: VARIABLE ESCARIFICADO MEDIANTE UN INDICADOR DE VECES.....	110
TABLA XXIX: TABLA DE FRECUENCIA DE ESCARIFICADO MEDIANTE UN INDICADOR DE VECES	111
TABLA XXX: ESTADÍSTICA DESCRIPTIVA: VARIABLE RIEGO	113
TABLA XXXI: TABLA DE FRECUENCIAS DE RIEGO	114
TABLA XXXII: ESTADÍSTICA DESCRIPTIVA: VARIABLE RIEGO MEDIANTE UN INDICADOR DE VECES	115
TABLA XXXIII: TABLA DE FRECUENCIAS RIEGO MEDIANTE UN INDICADOR DE VECES	116

TABLA XXXIV: ESTADÍSTICA DESCRIPTIVA: VARIABLE SACADA DE PAJA CAUCA	117
TABLA XXXV: TABLA DE FRECUENCIAS SACADA DE PAJA CAUCA	118
TABLA XXXVI: ESTADÍSTICA DESCRIPTIVA: VARIABLE SACADA DE PAJA CAUCA MEDIANTE UN INDICADOR DE VECES	120
TABLA XXXVII: TABLA DE FRECUENCIAS SACADA DE PAJA CAUCA MEDIANTE UN INDICADOR DE VECES.....	121
TABLA XXXVIII: ESTADÍSTICA DESCRIPTIVA: VARIABLE APORQUE	123
TABLA XXXIX : TABLA DE FRECUENCIAS DE APORQUE	124
TABLA XL: ESTADÍSTICA DESCRIPTIVA: VARIABLE APORQUE MEDIANTE UN INDICADOR DE VECES	125
TABLA XLI: TABLA DE FRECUENCIAS APORQUE MEDIANTE UN INDICADOR DE VECES	126
TABLA XLII: ESTADÍSTICA DESCRIPTIVA: VARIABLE FERTILIZACIÓN.....	128
TABLA XLIII: TABLA DE FRECUENCIAS DE FERTILIZACIÓN.....	129
TABLA XLIV: ESTADÍSTICA DESCRIPTIVA: VARIABLE FERTILIZACIÓN MEDIANTE UN INDICADOR DE VECES.....	130
TABLA XLV: TABLA DE FRECUENCIAS FERTILIZACIÓN MEDIANTE UN INDICADOR DE VECES	131
TABLA XLVI: NOMINA DE VARIABLES DE INTERÉS	134
TABLA XLVII: VALORES PROPIOS Y PROPORCIÓN DE LA VARIANZA TOTAL EXPLICADA DE LOS DATOS	135
TABLA XLVIII: VECTORES PROPIOS	136
TABLA XLIX: MATRIZ DE CORRELACIÓN DE LOS DATOS	138

TABLA L: PROPORCIÓN DE VARIANZA EXPLICADA (DATOS STANDARIZADOS).....	141
TABLA LI: VECTORES PROPIOS DE LOS DATOS STANDARIZADOS.....	143
TABLA LII: MATRIZ DE CARGA DE LOS DATOS STANDARIZADOS.....	145
TABLA LIII: DATOS ESTANDARIZADOS: PROPORCIÓN DE LA VARIANZA TOTAL EXPLICADA DE LOS FACTORES ROTADOS	147
TABLA LIV: MATRIZ DE CARGA DE LOS FACTORES ROTADOS.....	148

INDICE GRÁFICOS

GRAFICO I: SEMBRÍOS DE CAÑA DE AZÚCAR	23
GRAFICO II: PLANTA DE CAÑA DE AZÚCAR	25
GRAFICO III: COSECHA DE CAÑA	34
GRAFICO IV: USO DE LA CAPACIDAD INSTALADA DE AZUCAR.....	46
GRAFICO V: COMPORTAMIENTO DE LA SUPERFICIE SEMBRADA Y COSECHA.....	47
GRAFICO VI: PRECIOS AL PRODUCTOR DE CAÑA DE AZUCAR 1990-1998	50
GRAFICO VII: COMPORTAMIENTO DE LA PRODUCCION DE CAÑA DE AZUCAR.....	53
GRAFICO VIII: HISTOGRAMA DE FRECUENCIA	58
GRAFICO IX: OJIVA	59
GRAFICO X: DIAGRAMA DE CAJA	60
GRAFICO XI: RIEGO	79
GRAFICO XII: HISTOGRAMA DE FRECUENCIAS DE SUPERFICIE SEMBRADA...	83
GRAFICO XIII: OJIVA DE SUPERFICIE SEMBRADA.....	84
GRAFICO XIV: HISTOGRAMA DE FRECUENCIAS DE SUPERFICIE COSECHADA.....	87
GRAFICO XV: OJIVA DE SUPERFICIE COSECHADA.....	88
GRAFICO XVI: HISTOGRAMA DE FRECUENCIAS DE LA PROCCIÓN.....	90
GRAFICO XVII: OJIVA DE LA PRODUCCIÓN.....	91
GRAFICO XVIII: HISTOGRAMA DE FRECUENCIAS DE ROZA DE CANTERO	92
GRAFICO XIX: OJIVA DE ROZA DE CANTERO	94

GRAFICO XX: HISTOGRAMA DE FRECUENCIA DE ROZA DE CANTERO MEDIANTE UN INDICADOR DE VECES.....	96
GRAFICO XXI: OJIVA ROZA DE CANTERO MEDIANTE UN INDICADOR DE VECES	97
GRAFICO XXII: HISTOGRAMA DE FRECUENCIA DE HERBICIDAS.....	98
GRAFICO XXIII: OJIVA DE HERBICIDAS	99
GRAFICO XXIV: HISTOGRAMA DE FRECUENCIA DE HERBICIDAS MEDIANTE UN INDICADOR DE VECES	101
GRAFICO XXV: OJIVA DE HERBICIDAS MEDIANTE UN INDICADOR DE VECES	102
GRAFICO XXVI: HISTOGRAMA DE FRECUENCIA DE RASTRILLO	103
GRAFICO XXVII: OJIVA DE RASTRILLO.....	104
GRAFICO XXVIII: HISTOGRAMA DE FRECUENCIAS RASTRILLO MEDIANTE UN INDICADOR DE VECES	106
GRAFICO XXIX: OJIVA DE RASTRILLO MEDIANTE UN INDICADOR DE VECES	107
GRAFICO XXX: HISTOGRAMA DE FRECUENCIA DE ESCARIFICADO	108
GRAFICO XXXI: OJIVA DE ESCARIFICADO.....	109
GRAFICO XXXII: HISTOGRAMA DE FRECUENCIAS DE ESCARIFICADO MEDIANTE UN INDICADOR DE VECES.....	110
GRAFICO XXXIII: OJIVA DE ESCARIFICADO MEDIANTE UN INDICADOR DE VECES	112
GRAFICO XXXIV: HISTOGRA DE FRECUENCIA DE RIEGO.....	113
GRAFICO XXXV: OJIVA DE RIEGO.....	114
GRAFICO XXXVI: HISTOGRAMA DE FRECUENCIA RIEGO MEDIANTE UN INDICADOR DE VECES	115

GRAFICO XXXVII: OJIVA DE RIEGO MEDIANTE UN INDICADOR DE VECES	116
GRAFICO XXXVIII: HISTOGRAMA DE FRECUENCIAS SACADA DE PAJA CAUCA	117
GRAFICO XXXIX: OJIVA DE SACADA DE PAJA CAUCA.....	119
GRAFICO XL: HISTOGRAMA DE FRECUENCIAS DE SACADA DE PAJA CAUCA MEDIANTE UN INDICADOR DE VECES.....	120
GRAFICO XLI: OJIVA SACADA DE PAJA CAUCA MEDIANTE UN INDICADOR DE VECES	122
GRAFICO XLII: HISTOGRAMA DE FRECUENCIAS APORQUE.....	123
GRAFICO XLIII: OJIVA SACADA DE PAJA CAUCA MEDIANTE UN INDICADOR DE VECES	124
GRAFICO XLIV: HISTOGRAMA DE FRECUENCIAS APORQUE MEDIANTE UN INDICADOR DE VECES	125
GRAFICO XLV: OJIVA DE APORQUE MEDIANTE UN INDICADOR DE VECES....	127
GRAFICO XLVI: HISTOGRAMA DE FRECUENCIAS DE FERTILIZACIÓN	128
GRAFICO XLVII: OJIVA DE FERTILIZACIÓN	129
GRAFICO XLVIII: HISTOGRAMA DE FRECUENCIAS FERTILIZACIÓN MEDIANTE UN INDICADOR DE VECES.....	131
GRAFICO XLIX: OJIVA DE FERTILIZACIÓN MEDIANTE UN INDICADOR DE VECES	132
GRAFICO L: COMPONENTES PRINCIPALES	137
GRAFICO LI: VALORES PROPIOS DATOS ESTANDARIZADOS	142
GRAFICO LII: COMPONENTES EN EL PLANO	146

INTRODUCCIÓN

EL Cultivo y producción de la caña de azúcar, constituyen un sector relevante de la Economía de nuestro país, principalmente en la economía de algunas ciudades donde su principal ingreso es la producción de caña de azúcar, como es la ciudad de Milagro donde gran cantidad de personas trabajan en el Ingenio Valdez aportando con su mano de obra.

El objetivo de este estudio es determinar, e identificar las relaciones entre las diferentes variables que aparecen en el proceso de cultivo y producción de la caña de azúcar, además establecer grupos o asociaciones de estas variables para obtener nuevas variables ficticias o latentes que ayuden a entender como se esta desarrollando el proceso mencionado.

CAPITULO I

1. CULTIVO Y PRODUCCIÓN DE LA CAÑA DE AZÚCAR EN EL ECUADOR.



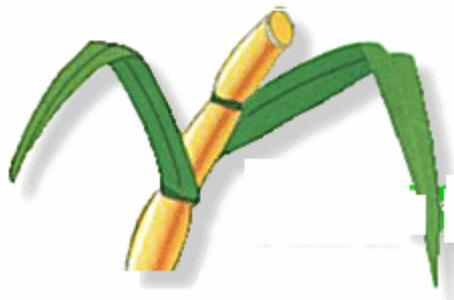
GRAFICO I: SEMBRÍOS DE CAÑA DE AZÚCAR

1.1 Generalidades de la caña de azúcar

La producción de caña en el Ecuador es realizada por 6 ingenios azucareros: La Troncal, San Carlos, Valdés, Isabel María, IANCEM y Monterrey, siendo los tres primeros quienes producen el 90 % de la producción nacional, estos ingenios conjuntamente con el Ingenio Isabel María se ubican en el Litoral Ecuatoriano, cuya zafra se inicia en el mes de julio y termina en diciembre, con procesos de molienda de 24 horas en tres turnos y un período interzafra (lo destinan exclusivamente a la reparación de maquinaria) entre enero - junio; Los ingenios IANCEM y Monterrey se encuentran localizados en la Región Sierra. La producción de azúcar se da en todo el año, trabajando seis días a la semana, el período interzafra lo realizan entre enero-febrero.

La producción de caña de azúcar es importante no solo por que por medio de ella obtenemos la azúcar sino también por que proporciona trabajo a un gran porcentaje de personas en la denominada zafra.

GRAFICO II: PLANTA DE CAÑA DE AZÚCAR



1.2 Antecedentes Históricos en el Ecuador

La caña de azúcar, caña dulce o cañaduz, fue cultivada y utilizada desde los tiempos más remotos.

El estudio de sus orígenes se hace muy difícil en la actualidad, debido a su amplia difusión y a la gran cantidad de cruces genéticos.

En la India, donde los soldados de Alejandro Magno la encontraron, en el 327 años A.C, durante su invasión a ese país. Uno de ellos escribía que los habitantes "masticaban una caña maravillosa que producía una especie de miel sin ayuda alguna de las abejas". No se sabe con certeza en que lugar se origino la caña que dio origen a la que se cultiva hoy. Los expertos presumen que se origino en alguna región de la India, o quizás de Nueva Guinea. De la India la planta paso a Mesopotamia, y de allí a Egipto, a través de las invasiones árabes, con las que se extendió por la cuenca del

Mediterráneo, a principios del siglo XIII. Se puede resumir que, desde Asia pasó al África, y de aquí a España, A partir de esta época, las plantaciones crecieron y en 1150 existía en España una floreciente industria azucarera.

La cultivaron en Valencia y Andalucía. La caña fue introducida en América por Cristóbal Colón, quien la llevo a Santo Domingo en su segundo viaje (1493), y plantada, por primera vez, en la isla La Española (República Dominicana).

De allí se expandió al resto de las islas del Caribe y otras regiones de América.

Este cultivo se desarrolló, entre los años 1500 y 1600, en la mayoría de los países tropicales de América (Las Antillas, México, Brasil, Perú, etc.) y durante mucho tiempo ha sido su principal riqueza agrícola.

En las nuevas regiones, la caña evolucionó, y adquirió nuevas características. En la zona de Polinesia, hasta Nueva Guinea, han sido colectadas diversas variedades de cañas que eran prácticamente cultivadas por los nativos. Al norte de la India, fue encontrado un grupo de cañas resistentes al frío (*Saccharum barberi*). Distribuido ampliamente, por esta región geográfica, crece el de cañas chinas (*Saccharum sinensi*) y los grupos de cañas silvestres (*Saccharum robustum* y *Saccharum spontaneum*), totalmente, resistentes a las enfermedades. A través del

cruzamiento de variedades pertenecientes a estos grupos, se han obtenido y obtienen, todos los años, nuevas variedades: las cañas híbridas.

La caña dulce pertenece a la familia de las gramíneas .Es una vigorosa planta perenne de crecimiento rápido que alcanza de tres a seis metros de altura y unos siete centímetros de diámetro.

Los tallos son nudosos, como los de toda caña, y salen de cepas robustas subterráneas formando haces. Las pequeñas flores se agrupan en grandes panojas plumosas en la terminación de los tallos, y constituyen un bello adorno de los campos. El tallo tiene una dura corteza fibrosa, como la del maíz, pero encierra un zumo azucarado que es la fuente más importante del azúcar comercial.

La caña dulce es una planta tropical. Requiere un clima húmedo y cálido favorecido con suficiente cantidad de lluvia. Las precipitaciones son más necesarias durante el crecimiento que en la época de maduración. Los cultivadores prefieren tiempo seco en la época de la cosecha, pues así la caña da un sumo mas concentrado .En cultivos de regadío, se puede dar a la caña la cantidad exactamente necesaria de agua para obtener un adecuado rendimiento.

En tiempos antiguos la caña dulce era planta de jardín o de huerta, pero actualmente se cultiva en grandes áreas “centrales” o ingenio, lo rodean casas para empleados, escuela, hospitales, tiendas y centros de recreo.

Generalmente se extienden alrededor de los ingenios azucareros inmensos campos de caña en todas direcciones. Los campos se aran cuidadosamente antes de la plantación de la caña. En vez de usar semilla, las plantaciones se hacen con estacas, que se entierran a lo largo de los surcos. En los últimos tiempos se ha reemplazado en muchas explotaciones la plantación a mano, la que se hace mediante maquinas. Según la región, la caña esta lista para la cosecha entre diez y treinta meses. Como las matas son perennes, las mismas sepas pueden producir varias cosechas, La caña se corta a machete, a flor de tierra, antes de florecer. Se conoce que esta bien madura por que se pone pesada y se rompe fácilmente.

No sólo se obtiene azúcar de la caña, sino también varios productos secundarios.

La caña triturada y exprimida constituye el bagazo, que puede servir como forraje o como abono, Seco, se lo utiliza también como combustible. En algunos países es materia prima para la fabricación de papel, cartón y sustancias plásticas. La melaza se emplea en la fabricación de alcohol, glicerina y varios ácidos, así como en la de diversos aguardientes.

La caña de azúcar es importante por que es la fuente de las tres cuartas partes de la producción mundial de azúcar.

1.3 FASE AGRICOLA

1.3.1 PROCESOS QUE SE NECESITAN PARA REALIZAR LA SIEMBRA DE CAÑA DE AZUCAR

Es recomendable, preparar el suelo dos meses antes de la siembra en época seca, el arado es de 30 a 40 centímetros de profundidad utilizando un arado de disco seguido de dos rastreadas, si se utiliza un arado de cinceles se debe hacer un pase de 20 a 25 centímetros de profundidad. En terrenos muy erosionados debe regularse para no provocar daños mayores. Durante la labor de preparación del suelo se recomienda aplicar 20 a 30 sacos de 110 libras cada uno de carbonato de calcio por hectárea, dependiendo del grado de acidez reportado en el análisis del suelo.

Se consulta siempre la mejor distancia para la variedad a sembrar entre una y otra planta con ayuda de los técnicos ya que ellos ayudaran a escoger la mejor opción.

1.3.1.2 SEMILLEROS DE CAÑA

Los semilleros son establecidos con dos fines :

- Asegurar la pureza y sanidad de la variedad a reproducir.

- Obtener semilla en condiciones óptimas de germinación y desarrollo vegetativo.

La semilla de caña es el material vegetativo de siembra, que debe tener entre 7 y 9 meses de edad y que propaga en trozos de 35-40 centímetros de longitud (3 yemas). El material usado como semilla debe ser puro en cuanto a la variedad, vigoroso en su germinación y libre de plagas y enfermedades.

Una semilla de calidad es obtenida de una buena cosecha, es por ello que se deben establecer semilleros en lotes dedicados para este fin y que reciban un manejo tal que le permita al productor garantizar la calidad del material a obtener.

Después de los nueve meses, el poder germinativo del material se reduce, haciéndose necesario sembrar más cantidad de cañas por hectárea.

El área del semillero que se debe establecer debe responder a las necesidades de la plantación en cuánto al área a sembrar por año.

Un semillero proporciona

Semilla en relación 1:6 o sea que si un productor va a sembrar 6 ha entre renovación y áreas nuevas debe establecer un semillero de 1 ha para suplir las necesidades de semilla.

A continuación vamos a presentar las necesidades de metros de semilla por hectárea.

TABLA I: SEMILLEROS POR HECTAREA

AREA A SEMBRAR(ha)	AREA SEMILLERO(m)
12	20.000
9	15.000
6	10.000
3	5.000
2	3.333
1	1.667

Los aspectos que deben considerarse al realizar un semillero de caña son :

- El área del semillero en función del área a sembrar
- Debe sembrarse de manera que la cosecha coincida con una edad de 7-9 meses (siembra en setiembre para cosechar en mayo).
- Debe ubicarse estratégicamente en la finca que le permita reducir gastos de transporte y que sea accesible a riego.
- No se debe realizar más de tres cortes en un mismo semillero. A partir del primer corte, el rebrote se debe cortar en septiembre para ajustar la edad de la semilla a 9 meses.
- Debe utilizarse un plan de fertilización basado en:
 1. Abonado : 12 sacos en 10/ha al fondo del surco

2. Abonado : 6 sacos Nutran/ha a los 30 días

3. Abonado : 8 sacos de Nutramon/ha

- La fertilización puede ser utilizada según los niveles nutricionales; para el caso del segundo y tercer corte se puede utilizar segunda y tercera abonada.

1.3.1.3 SELECCIÓN DE LA SEMILLA

Para el proceso de siembra es muy importante seleccionar la semilla apropiada ya que dependerá de la misma para obtener una buena cosecha.

Entre las variedades de caña de azúcar tenemos:

La Variedad SP-71-5574.- Obtiene su máximo de madurez en caña planta y caña soca entre 11 y 12 meses, esta variedad a los 11 meses ya tiene un concentrado de sacarosa, manteniendo hasta los 13 meses lo que amplía su periodo de cosecha aspecto muy importante ya que se programa ella misma a su cosecha.

La Variedad Pindar .- Obtiene su máximo de madurez en caña planta a los 11 meses y en soca a los 11.5 meses, es una variedad que mantiene poco tiempo la concentración de azúcar y rápidamente

declina. Por ello se debe realizar la cosecha en el preciso momento de su madurez fisiológica para que no se vea afectado el rendimiento de la cosecha.

La Variedad B-47-44.- Obtiene su máximo de madurez en caña planta a los 11 meses y en soca a los 12 meses, esta es una variedad que tiene menor concentración de azúcar que las anteriores mencionadas.

1.3.2 SIEMBRA

Luego de seleccionar la caña que se desea sembrar procedemos a la siembra de la misma de una manera rápida para evitar brotamientos o deterioro de las yemas antes de sembrarse.

Es necesario picar los tallos de la semilla en trozos de 3 yemas cada uno, Se debe realizar una adecuada fertilización fosfórica aplicada al fondo del surco, si cuenta con un análisis químico del suelo debe ajustarse a los resultados del mismo.

Finalmente se debe tapar la semilla tan pronto como sea posible con una capa de tierra no mayor a 10 cm ,debemos recordar que se realice por lo menos dos semanas antes el calado de la tierra en la cual se va a efectuar la siembra de la semilla ,es importante que antes de echar la cal al suelo el suelo se encuentre húmedo, se recomienda extender la cal por todo el área de la plantación no solo en el surco .El brotamiento

de la misma es 5 – 30 días después de la siembra (caña planta) inmediatamente después del corte (caña soca)

1.3.3 ETAPAS DE LA CAÑA DE AZUCAR DURANTE SU RECIMIENTO

Primera Etapa.

En la primera etapa se la denomina “caña planta” que alcanza una altura de 30 centímetro aproximadamente se la denomina así hasta que le realicen su primer corte, tiene un ciclo de vida de 14 – 18 meses , donde siempre se realiza la primera cosecha

Segunda Etapa.

En la segunda etapa se la denomina “caña soca” después de haberle echo su primer corte, en su primera cosecha.

Esta tiene un tamaño de de 2 a 4 metros por año, Diámetro de 2 a 6 centímetros.

1.3.4 COSECHA

GRAFICO III: COSECHA DE CAÑA



La cosecha se realiza en caña planta a los 21 meses, caña soca entre los 15 y 18 meses.

Luego de que la planta se encuentre madura las personas encargadas del área de cosecha se disponen a cortarla y recogerla a través de alce mecánico y llevarla hacia los patios de caña de los ingenios.

PATIOS DE CAÑA

La caña recolectada del campo la cual es enviada a una superficie vacía mediante el cual se determina las características de calidad y que cantidad de sacarosa, fibra y nivel de impurezas poseen. Luego es trasladada en básculas y se conduce a los patios donde se almacena temporalmente o se dispone directamente en las mesas de lavado de caña para dirigirla a una banda conductora que alimenta las picadoras.

PICADO DE CAÑA

Las picadoras son unos ejes colocados sobre los conductores accionados por turbinas, provistos de cuchillas giradoras que cortan los tallos y los convierten en astillas, dándoles un tamaño uniforme para facilitar así la extracción del jugo en los molinos.

MOLIENDA

Luego de haber preparado la caña por las picadoras llega a un tándem de molinos, constituido cada uno de ellos por tres o cuatro mazas metálicas y mediante presión extrae el jugo de la caña. Cada molino esta equipado con una turbina de alta presión. En el recorrido de la caña por el molino se agrega agua, generalmente caliente, para extraer al máximo la sacarosa que contiene el material fibroso. Este proceso de extracción es llamado maceración. El bagazo que sale de la última unidad de molienda se conduce a una bagacera para que seque y luego se va a las calderas como combustible, produciendo el vapor de alta presión que se emplea en las turbinas de los molinos.

PESADO DE JUGOS

El jugo diluido que se extrae de la molienda se pesa en básculas con celdas de carga para saber la cantidad de jugo sacarosa que entra en la fábrica.

CLARIFICACION

El jugo obtenido en la etapa de molienda es de carácter ácido (pH aproximado: 5,2), éste se trata con lechada de cal, la cual eleva el pH con el objetivo de minimizar las posibles pérdidas de sacarosa. La cal también ayuda a precipitar impurezas orgánicas o inorgánicas que vienen en el jugo y para aumentar o acelerar su poder coagulante, se eleva la temperatura del jugo encalado mediante un sistema de tubos calentadores. La clarificación del jugo por sedimentación, los sólidos se precipitan en forma de lodo llamado cachaza y el jugo claro queda en la parte superior del tanque. Este jugo sobrante se envía antes de ser desechada al campo para el mejoramiento de los suelos pobres en materia orgánica.

EVAPORACIÓN

Donde se comienza a evaporar el agua del jugo. El jugo claro que posee casi la mitad composición del jugo crudo extraído (con la excepción de las impurezas eliminadas en la cachaza).

Se recibe en los evaporadores con un porcentaje de sólidos solubles entre 10 y 12 % y se obtiene una meladura o jarabe con una concentración aproximada de sólidos solubles del 55 al 60%.

Este proceso se da en evaporadores de múltiples efectos al vacío; que consisten en una solución de celdas de ebullición dispuestas en serie. El jugo entra primero en el preevaporador y se calienta hasta el punto de ebullición. Al comenzar a ebullición se generan vapores los cuales sirven para calentar el jugo en el siguiente efecto, logrando así al menor punto de ebullición en cada evaporador. En el proceso de evaporación se obtiene el jarabe o meladura. La meladura es purificada en un clarificador. La operación es similar a la anterior para clarificar el jugo filtrado.

CRISTALIZACIÓN

La cristalización se realiza en los tachos, que son recipientes al vacío. El material resultante que contiene líquido (miel) y cristales (azúcar) se denomina masa cocida. El trabajo de cristalización se lleva a cabo empleando el sistema de tres cocimientos para lograr la mayor concentración de sacarosa.

CENTRIFUGACIÓN

La masa pasa por las centrífugas, máquinas agrícolas en las cuales los cristales se separan del licor madre por medio de una masa centrífuga aplicada a tambores rotatorios que contienen mallas interiores. La miel que sale de las centrifugas se bombea a tanques de almacenamiento para

luego someterla a otras evaporaciones y cristalizaciones en los tachos. Al cabo de tres cristalizaciones sucesivas se obtiene miel final que se retira del proceso y se comercializa como materia prima para la elaboración de alcoholes.

SECADO

El azúcar húmedo se transporta por elevadores y bandas para alimentar las secadoras que son elevadores rotatorios en los cuales el azúcar se colocan en contacto con el aire caliente que entra en contracorriente. El azúcar debe tener baja humedad, aproximadamente 0.05 %, para evitar los terrones.

ENFRIAMIENTO

El azúcar se seca con temperatura cercana a 60°C, se pasa por los enfriadores rotatorios inclinados que llevan el aire frío en contracorriente, en donde se disminuye su temperatura hasta aproximadamente 40-45°C para conducir al envase.

ENVASE

El azúcar seca y fría se empaca en sacos de diferentes pesos y presentaciones dependiendo del mercado y se despacha a la bodega de producto terminado para su posterior venta y comercio.

1.3.5 PRINCIPALES ENFERMEDADES, PLAGAS Y MALEZAS DE LA CAÑA DE AZÚCAR.

Las enfermedades son un factor muy importante ya que depende mucho de ellas el debido desarrollo del crecimiento de la planta lo cual relaciona directamente el éxito de la cosecha de la caña de azúcar.

1.3.5.1 ENFERMEDADES

TABLA II: ENFERMEDADES DE LA CAÑA DE AZÚCAR

LOCALIZACION	NOMBRE COMUN	TIPO	SINTOMA
Hoja	Roya	Hongo	Pústulas color errumbre
Hoja y Brote	Polvillo	Bacteria	Rayas rojas y muerte brote guía
Hoja y tallo	Escaldadura	Bacteria	Estrías blancas y muerte de tallo
Hoja	Mosaico	Virus	Clorosis de hoja
Brote	Carbón	Hongo	Látigo y polvillo de carbón
Brote y tallo	Raquitismo	Virus	Enanismo de cepas de más de 1 año

1.3.5.2 PLAGAS

Entre las plagas tenemos:

El Joboto.- constituyen una plaga del suelo que causa daños severos en los rendimientos de los cultivos y entre estos cultivos se encuentra la caña de azúcar.

Este insecto provoca una poda de raíz en la caña. Causando que la caña se quede raquílica provocando un gran daño en su crecimiento.

Esto no cambia por más fertilizante que le apliquemos debido que la ubicación estratégica que adquiere el joboto dentro del suelo, debajo de la cepa utilizando esta como escudo protector. Su habilidad de profundizarse en el suelo, hasta puntos donde el efecto de los químicos es difícil que llegue.

Se recomienda estrategia de control como:

Trampas de luz: Aprovechar la característica de los abejones de ser atraídos por la luz, para ser atrapados y eliminados.

Gusano Talador

Esta plaga produce daño a la caña de azúcar atacando en tallos lo cual perjudica a su crecimiento y a su desarrollo normal.

También tenemos otras plagas como el Salivazo y Baba de Culebra

1.3.5.3 MALEZAS

Las malezas compiten por agua, luz y lo más grave por nutrimento con la caña en muchos casos son hospederos de plagas y enfermedades. En el cultivo de caña el control se hace necesario en los primeros tres meses de establecido el cañal o después de cortado en la denominada caña soca.

Debemos considerar algunos aspectos para el control de malezas utilizando herbicidas tales como:

- El herbicida no debe sobrepasar los 10 centímetros de altura.
- El suelo debe estar húmedo. No se debe aplicar cuando va a llover o hace viento fuerte.
- Utilizar agua limpia.

Entre las malezas tenemos:

- Grama Bermuda
- Pasto Ruso
- Cebollin
- camalote

2. CULTIVO DE LA CAÑA DE AZÚCAR EN EL ECUADOR

2.1 SUPERFICIE SEMBRADA

Se ha ido incrementando la superficie sembrada al transcurrir el tiempo, tiene un crecimiento sostenido como podemos darnos cuenta en 1990 se sembraron 48.201 Has. de las cuales se cosecharon 45.642 Has. Quienes producen 3.391,525 TM. , pasando a 1998, donde nos representa un incremento del 40% ,y de igual manera se observa en el área cosechada ,cuya excepción es la de 1997 , que fue afectada por la presencia del Fenómeno del Niño ,en la cual fueron cosechadas 24.463 Has de caña de azúcar , lo cual produjo una reducción de la superficie del 57% con respecto a 1996.

TABLA III: PRODUCCIÓN DE AZÚCAR DEL ECUADOR

Años 1/	Superficie Sembrada (ha)	Superficie por Cosechar (ha)	Producción Caña (TM)	Producción Azúcar		Rendimientos	
				(TM)	Sacos 50 Kg.	TM Caña/ha	Lb. Azú/TMc
1990	48,201	45,642	3,391,525	331,925	6,638,497	74.31	216
1991	50,264	48,200	3,612,678	325,656	6,513,124	74.95	199
1992	50,248	43,628	3,757,514	358,285	7,165,702	86.13	210
1993	54,011	49,893	3,666,270	338,031	6,760,620	73.48	203
1994	54,061	49,516	3,398,428	319,970	6,399,394	68.63	208

1995	56,793	53,280	3,895,744	364,923	7,298,469	73.12	207
1996	57.851	57.471	4.412.519	437.730	8.754.598	76,78	219
1997/2	69732	24.463	2.527.215	186.262	3.725.237	70,77	162
1998	67.469	67.403	4.470.457	371.688	6.741.391	66,32	157
1999	68.000	67.240	4.529.238	393.946	7.878.916	67,36	170
2000	70.000	69.085	4.662.322	453.750	9.075.000	72,00	193
2001*	70.172	66.462	4.662.322	453.750	9.075.000	72,00	193

En el Ecuador existen 6 ingenios de azúcar mediante el cuadro siguiente mostraremos la superficie sembrada por cada uno de ellos.

**TABLA IV: HECTAREA SEMBRADA DE LOS INGENIOS DEL
ECUADOR DE Junio/2000 -Junio /2001**

INGENIOS	HECTÁREAS	HECTÁREAS COSECHADAS			RENDIT	PRODUCCIÓN
	SEMBRADAS	INGENIO	CAÑICULTOR	TOTAL	TMC/HAS	TM-CAÑA
INGENIO VALDEZ	16.340	10.270	6.070	16.340	66	1.085.592
INGENIO SAN CARLOS	20.400	7.666	12.734	20.400	75	1.528.731
INGENIO LA TRONCAL	26.520	9.603	16.917	26.520	61	1.612.893
INGENIO MONTERREY	1.653	875	778	1.653	101	167.064
INGENIO IANCEM	3.672	612	3.060	3.672	61	225.542
INGENIO ISABEL MARIA	500	-	500	500	70	42.500
TOTAL	69.085	29.026	40.059	69.085	72	4.662.322

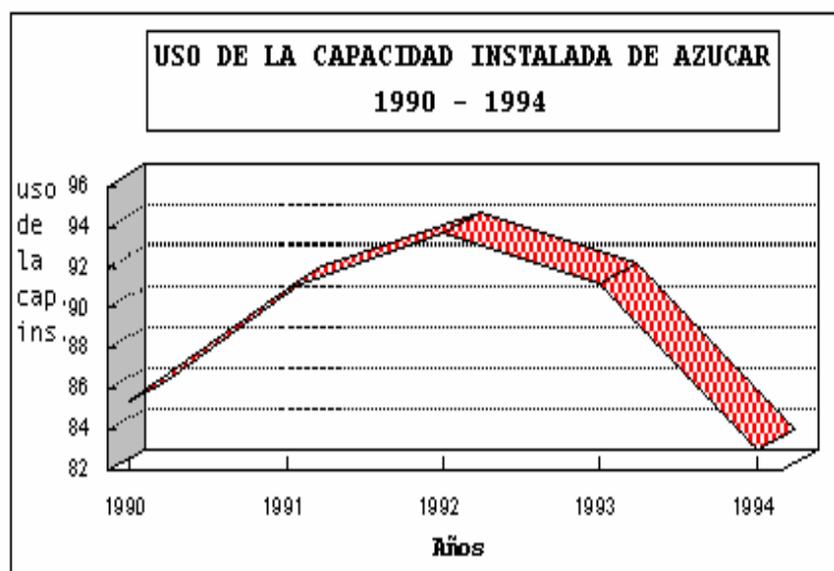
2.2 CAPACIDAD INSTALADA POR INGENIO

La mejor forma de medir la productividad en fábrica es mediante la proporción en que se utiliza la capacidad instalada de molienda de los ingenios azucareros.

En el gráfico IV, se puede observar el uso de la capacidad instalada de los ingenios azucareros en el período 1990 - 1994.

En el período de análisis se evidencia que la productividad en fábrica varió dentro del rango de 83 % a 94 %, el año de mayor utilización de la capacidad instalada fue en 1993 con 94%.

La variación en el uso de la capacidad instalada de los ingenios esta en función prioritaria del abastecimiento de materia prima que pueda conseguir. Por este motivo, futuros incrementos en el uso de la capacidad instalada dependerán de que aumente la producción de caña en el país, ya sea mediante mayores rendimientos en el campo o a través de una mayor superficie cultivada.

GRAFICO IV: USO DE LA CAPACIDAD INSTALADA DE AZUCAR

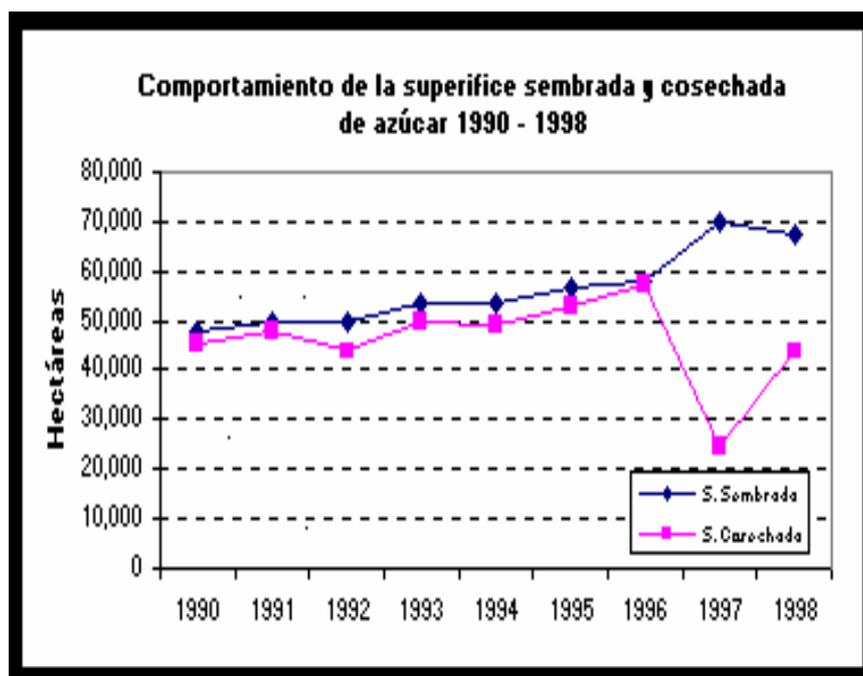
A partir de 1996, los principales ingenios azucareros han invertido parte de sus utilidades en la ampliación de su capacidad de molienda, es así como la producción de azúcar obtenida en 1996 rompió todos los récords de producción del país.

El volumen de la caña cosechada es igual al de la caña molida

TABLA V: CAPACIDAD INSTALADA

AÑOS	CAPACIDAD INSTALADA TM/día 1]	CAPACIDAD UTILIZADA TM/día 2]	% DE USO
1990	26,5	22,61	85.32
1991	26,5	24,085	90.89
1992	26,75	25,05	93.64
1993	26,75	24,386	91.16
1994	26,75	22,196	82.98

GRAFICO V: COMPORTAMIENTO DE LA SUPERFICIE
SEMBRADA Y COSECHA



2.2.1 PRECIOS A NIVEL DE FINCA

Hasta el año 1992, los precios a nivel de finca eran fijados por el Gobierno Nacional, basándose en los costos de producción.

Es a partir de enero de 1993, que como consecuencia de la política de liberalización, los precios de caña de azúcar, azúcar y melaza, se rigen aparentemente de acuerdo a la oferta Y demanda, es decir, al libre juego del mercado.

Para el primer semestre de 1993, los precios de la caña de azúcar fueron de S/. 16.466 cada T.M. en pie con 13 grados de sacarosa, más un premio adicional por décima de gramo superior a los 13 grados de S/. 54.3.

Posteriormente para el segundo semestre/93, entró a regir el precio promedio de S/. 18.871 (precio concertado), lo que significó un incremento del 8.3.

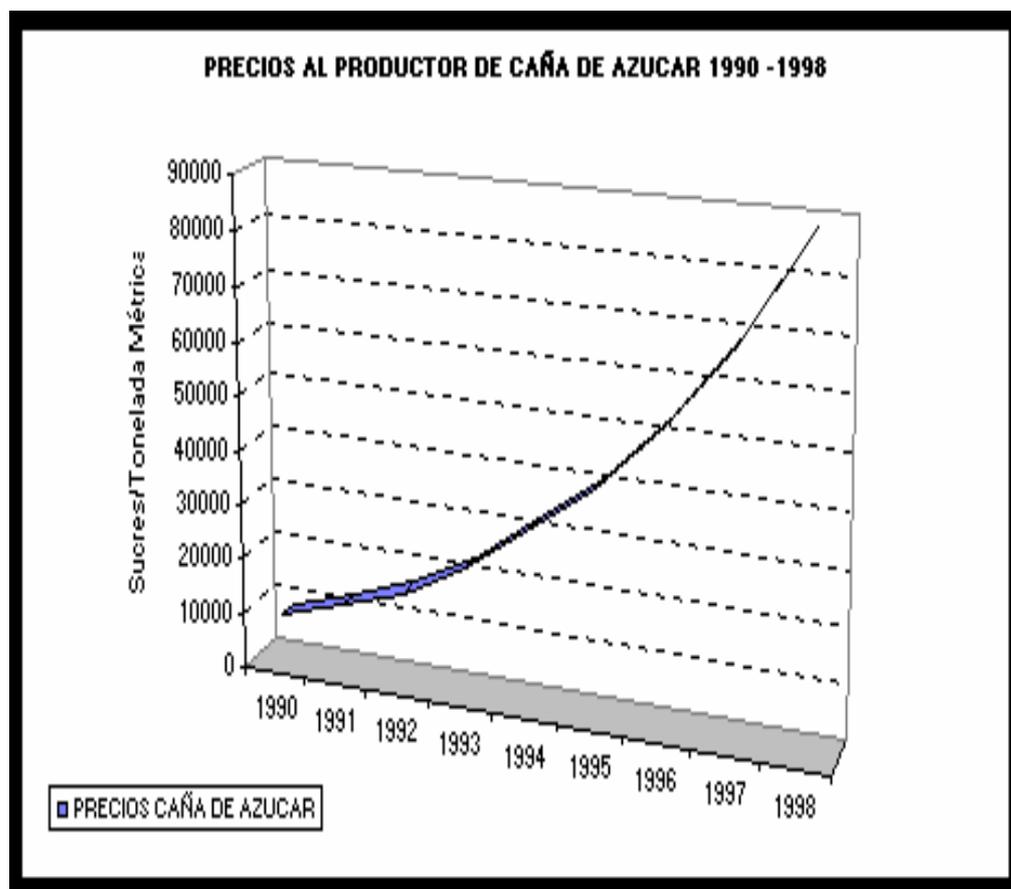
En la zafra julio/93 - junio/94 el precio que los agricultores recibieron por parte de los ingenios fue equivalente al 75% del valor del precio de venta del saco de 50 Kgrs de azúcar ex-ingenio.

**TABLA VI: PRECIOS NOMINALES DE LA CAÑA DE AZUCAR
DE ENERO A OCTUBRE DE 1998**

MESES	(sucres/M)
ENERO	73,5
FEBRERO	73,5
MARZO	76
ABRIL	82
MAYO	88,25
JUNIO	89,25
JULIO	92,75
AGOSTO	93,75
SEPTIEMBRE	106,75
OCTUBRE	106,75
NOVIEMBRE	111
DICIEMBRE	111
PROMEDIO	88,25
MINIMO	73,5
MAXIMO	106,75

En 1998 el precio de TM. Fue en promedio de 88.250 sucres, lo que representa un incremento nominal del 90 % en relación al precio registrado en 1990 que fue de 46.437 sucres por tonelada métrica de caña.

**GRAFICO VI: PRECIOS AL PRODUCTOR DE CAÑA DE AZUCAR
1990-1998**



El precio máximo registrado en 1998, fue de 111.000 s/. Por TM, el mismo que se observó en el mes de diciembre/98, precio que representó un estímulo para los cañicultores, quienes tuvieron una mayor rentabilidad, situación que se vio reflejada en el incremento de la producción para el ciclo 1998-1999.

TABLA VII: PRECIOS PARA LOS CAÑICULTORES Años 1999-2000 (US\$/TM)

AÑO 2000	
ENERO	12,75
FEBRERO	12,75
MARZO	13,5
ABRIL	13,5
MAYO	13,5
JUNIO	13,5
AÑO 1999	
ENERO	15,57
FEBRERO	15,92
MARZO	15,71
ABRIL	16,34
MAYO	17,06
JUNIO	16,47
JULIO	15,01
AGOSTO	14,78
SEPTIEMBRE	14,28
OCTUBRE	12
NOVIEMBRE	13,12
DICIEMBRE	12,69
Fuente: FENAZUCAR	

2.3 PRODUCCION DE LA CAÑA DE AZÚCAR

Como se puede apreciar en el Gráfico VII, la producción de caña de azúcar ha tenido un crecimiento constante excepto en el año de 1997, que se obtuvo decrecimiento por la presencia del fenómeno del niño en un 25% con relación a 1990 y del 43% con respecto a 1996, la producción de caña

1998, fue de 4`986.745 TM., lo que equivale a un crecimiento de alrededor del 47% con respecto a 1990.

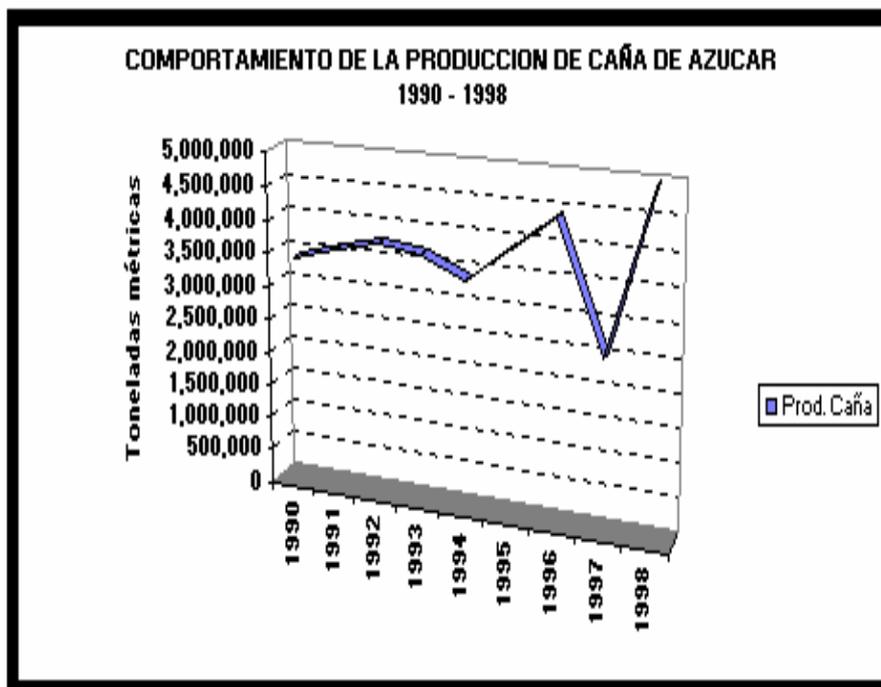
La baja de la producción de caña en 1997 se debió a problemas climáticos, que se presentaron dejando de zafrar una área de 42.444 Ha., las mismas que si fueron cosechadas en 1998, año en el cual además los ingenios azucareros y los cañicultores, realizaron innovaciones tecnológicas lo cual lográndose exitosamente la productividad vía rendimiento .

Otro factor que incidió en el crecimiento de la producción en el periodo de análisis fue la política de liberación de precios que se aplicó a partir del año 1993, medida que representó un estímulo para los cañicultores, que significaba un precio satisfactorio que justificaba sus costos de producción y además les permitía obtener ganancias.

Entre 1990 y 1996 el mayor rendimiento industrial fue obtenido en el último año, cuando por el intenso verano, se incrementó el contenido de sacarosa en la caña, lo cual permitió obtener un alto rendimiento de 224 libras por Toneladas métricas de caña, y el rendimiento menor se ubicó en 1993 y 1997, debido a las inundaciones registradas especialmente en el Litoral Ecuatoriano.

En efecto, como consecuencia del invierno, no solo dificultó el proceso de cosecha, causando también la reducción de sacarosa en la planta, provocando la disminución del rendimiento, lo que trajo consigo una mala calidad de caña.

**GRAFICO VII: COMPORTAMIENTO DE LA PRODUCCION DE
CAÑA DE AZUCAR**



CAPITULO II

DESCRIPCION DE LAS TECNICAS ESTADISTICAS A UTILIZAR

INTRODUCCIÓN

Para realizar un análisis estadístico es recomendable realizar un análisis univariado, incluyendo técnicas descriptivas e inferenciales.

2.1 Análisis univariado

Mediante una muestra aleatoria de observaciones seleccionada de una población, se obtienen la media aritmética y la varianza muestral, además se elaboran gráficos representativos como: Función empírica, histograma de frecuencia, ojiva y diagrama de cajas.

La media aritmética y la varianza muestral son fundamentales para describir el comportamiento de los datos de una muestra aleatoria dada, de allí que sea importante conocer sus respectivas definiciones.

Sean x_1, x_2, \dots, x_n , elementos de una muestra aleatoria de tamaño n , la media aritmética, denotada por \bar{x} , esta definida por:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

Sean x_1, x_2, \dots, x_n , elementos de una muestra aleatoria de tamaño n , y sea \bar{x} la media aritmética de la muestra, la varianza muestral, denotada por s^2 , se define de la siguiente forma:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Una representación importante en un análisis estadístico son los gráficos, en cuanto a su dispersión, tendencia central, frecuencias relativas y acumuladas. A continuación se definirán los gráficos mas importantes:

Función empírica, histograma de frecuencias, ojiva y diagrama de cajas.

Sean $x_{(1)}, x_{(2)}, \dots, x_{(k)}, \dots, x_{(n)}$, los estadísticos de orden correspondiente a una muestra aleatoria de tamaño n , esto es $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$, la función empírica, denotada por $\hat{F}(X)$ se define de la siguiente forma:

$$x_{(1)} = \min(x_1, x_2, \dots, x_n)$$

$x_{(k)}$ es el k -ésimo estadístico de orden ; es el k -ésimo estadístico de orden;

$$2 \leq k \leq n - 1$$

$$x_{(n)} = \max(x_1, x_2, \dots, x_n)$$

$$\hat{F}(x) = \begin{cases} 0 & ; x < x_{(1)} \\ \frac{k}{n} & ; x_{(k)} \leq x < x_{(k+1)} \\ 1 & ; x \geq x_{(n)} \end{cases}$$

para dar una mejor ilustración presentaremos un ejemplo.

Si tenemos la siguiente muestra de tamaño 6 que es :

$x_1 = 4, x_2 = 2, x_3 = 1, x_4 = 5, x_5 = 6, x_6 = 7$ obtendremos los siguientes estadísticos de orden:

$x_{(1)} = 1, x_{(2)} = 2, x_{(3)} = 4, x_{(4)} = 5, x_{(5)} = 6, x_{(6)} = 7$ obteniendo como resultado la siguiente función empírica:

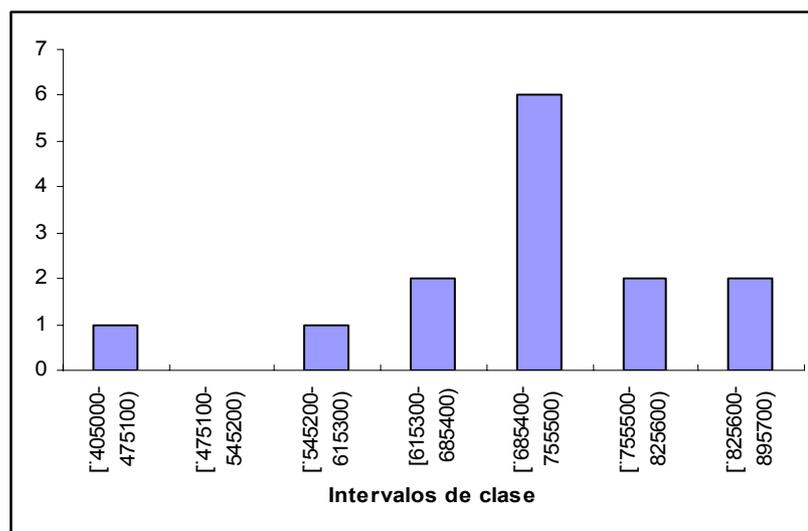
$$\hat{F}(X) = \begin{cases} 0; X < 1 \\ \frac{1}{6}; 1 \leq X < 2 \\ \frac{2}{6}; 2 \leq X < 4 \\ \frac{3}{6}; 4 \leq X < 5 \\ \frac{4}{6}; 5 \leq X < 6 \\ \frac{5}{6}; 6 \leq X < 7 \\ 1; X \geq 7 \end{cases}$$

2.1.1 HISTOGRAMA DE FRECUENCIAS

Las clases deben ser de igual longitud de tal forma que la primera clase incluya al mínimo valor $x_{(1)}$ y la última incluya al máximo valor $x_{(n)}$, además son intervalos consecutivos.

Un histograma de frecuencias está formado por una serie de barras que tienen sus bases sobre un eje horizontal (eje x) e iguales al ancho de su clase y su altura es igual a la frecuencia relativa de clase.

GRAFICO VIII: HISTOGRAMA DE FRECUENCIA



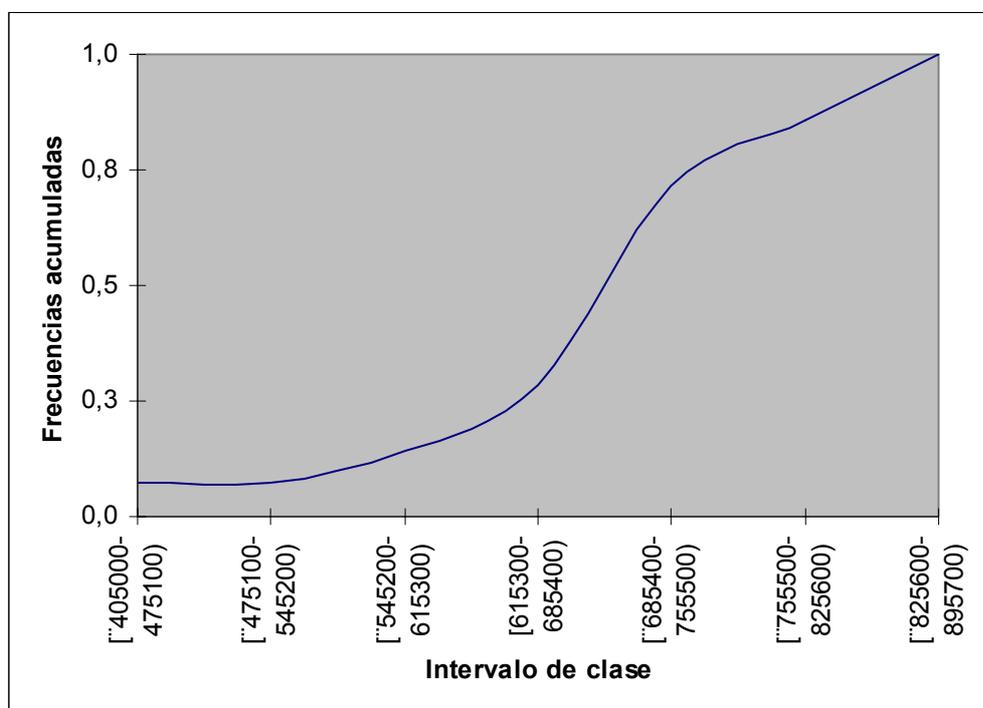
2.1.2 OJIVA

Es el grafico de líneas trazadas sobre las clases definidas para el histograma de frecuencias.

La ojiva tiene una frecuencia acumulada relativa igual a cero en el mínimo valor de la primera clase e igual a uno en el máximo valor de la ultima clase. En la grafica de ojiva se pueden determinar los percentiles p_1, p_2, \dots, p_{99} , donde el 1% de las observaciones son menores o iguales que p_1 , el 2% de las observaciones son menores

o iguales que p_2 , y así sucesivamente; los deciles D_1, D_2, \dots, D_9 , donde el 10% de las observaciones son menores o iguales que D_1 , el 20% de las observaciones son menores o iguales que D_2 , y así sucesivamente; y los cuarteles Q_1, Q_2 Y Q_3 donde el 25% de las observaciones son menores o iguales que Q_1 , el 50% de las observaciones son menores o iguales que Q_2 Y el 75% de las observaciones son menores o iguales que Q_3 .

GRAFICO IX: OJIVA

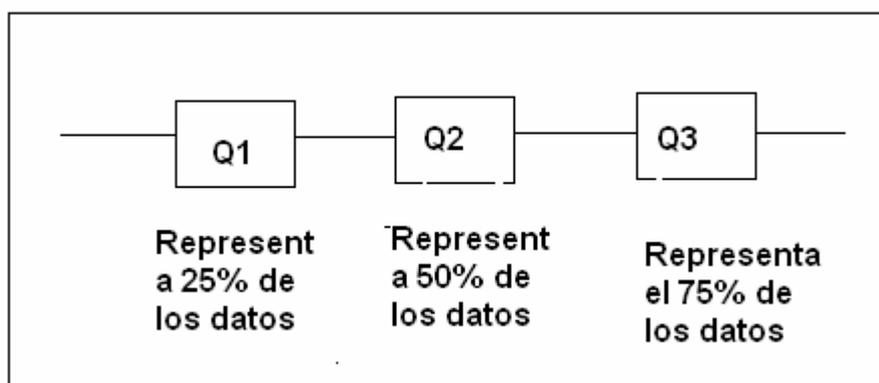


2.1.3 DIAGRAMA DE CAJAS

Un diagrama de cajas nos puede mostrar la tendencia central de los datos y a su vez la dispersión de los mismos. Su grafico nos puede representar valores muy grandes los cuales se representan mediante asteriscos (*).

El diagrama también nos representa los cuarteles Q_1 , Q_2 y Q_3 y dos segmentos de recta, la una que es la línea que se traza desde el mínimo valor hasta el primer cuartil y el segundo que es la línea que se traza desde el tercer cuartil hasta el valor máximo. el valor de Q_2 representa la mediana muestral.

GRAFICO X: DIAGRAMA DE CAJA



2.1.4 PRUEBA DE BONDAD DE AJUSTE: KOLMOGOROV-SMIRNOV

La prueba de Kolmogorov-Smirnov es una prueba de bondad de ajuste.

Esta prueba permite determinar si la distribución propuesta es un modelo satisfactorio para la población estudiada, evaluando la concordancia entre los valores de la distribución empírica de la muestra y la función de distribución especificada. La prueba se basa en la máxima diferencia absoluta D entre los valores de la distribución acumulada empírica de una muestra aleatoria de tamaño n y una distribución acumulada teórica específica.

La prueba de hipótesis se establece así:

H_0 : La población muestreada tiene función de distribución $F_0(X)$

vs.

H_1 : No es verdadero H_0

2.2 ANALISIS MULTIVARIADO

En el análisis multivariado se utilizan varias maneras para representar el universo de estudio tales como la simplificación de la estructura de datos, mediante la transformación (combinación lineal o no lineal) de un conjunto de variables independiente o en un conjunto de menor dimensión.

Este tipo de análisis permite ubicar las observaciones dentro de grupos o bien concluir que los individuos están dispersos aleatoriamente en el multiespacio; también se puede agrupar variables.

El objetivo es analizar la interdependencia de las variables, la cual cubre desde la independencia total hasta la colinealidad cuando una de ellas es combinación lineal de alguna de las otras o, generalizando es una función $F(x)$ cualquiera de las otras.

Entre los varios métodos de análisis multivariado para detectar la interdependencia entre variables y también entre individuos se incluyen el análisis de factores, el análisis por conglomerado, el análisis de correlación canónica, el análisis por componentes principales, el análisis de ordenamiento multidimensional, y algunos métodos no paramétricos.

Los métodos para detectar dependencia son: el análisis de regresión multivariado, el análisis de contingencia múltiple y el análisis discriminante.

Uno de los métodos más utilizados es el método de análisis de componentes principales, el cual permite la estructuración de un conjunto de datos multivariados obtenidos de una población.

2.2.1 DEFINICIONES

Para interpretar los resultados de un análisis de componentes principales se necesita tener presente un conjunto de conceptos, que son de mucha importancia, los cuales presentaremos a continuación.

2.2.2 MATRIZ DE DATOS

Decimos que un grupo de datos constituye una muestra aleatoria multivariada si cada individuo ha sido extraído al azar de una población de individuos y en él se han medido u observado muchas características. Una vez definido muestra aleatoria multivariante, ahora trataremos lo que se conoce como matriz aleatoria. Matriz aleatoria es una matriz cuyos elementos son variables aleatorias por p variables aleatorias o características $p \geq 1$ y n observaciones

registrada para cada variable. Sean $X(ij)$ la observación de la j -ésima variable en el i -ésimo individuo, $X(i)$ el vector fila que contiene las observaciones de todas las variables en el i -ésimo individuo y $X(j)$ el vector columna que contiene todas las observaciones de la j -ésima variable. Se define una matriz de datos como el arreglo de:

$$X = (X(ij)) = \begin{bmatrix} X(11) & \dots & \dots & X(1p) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & X(ij) & \vdots \\ X(n1) & \vdots & \vdots & X(np) \end{bmatrix}$$

de dimensión $n \times p$ que también puede expresarse como

$$X = \langle X(1) \dots X(p) \rangle = \begin{bmatrix} X(1) \\ \vdots \\ \vdots \\ X(n) \end{bmatrix}$$

2.2.3 VECTOR DE MEDIAS Y MATRICES DE VARIANZA Y COVARIANZAS

La media muestral de la j -ésima variable de una matriz de datos se define por:

$$E(X_j) = 1/n \sum_{i=1}^n X_{ij}$$

El vector formado por los $E(X_j)$ será el vector de medias .

$$\mu = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix}$$

dada una matriz de datos, la varianza muestral de la j -ésima variable se la define por:

$$S(j) = 1/n \sum_1^n (X_{ij} - \bar{X}(j))^2$$

y se define la covarianza entre la j -ésima y la k -ésima variable por:

$$S_{ik} = 1/n \sum_1^n (X_{ij} - \bar{X}(j))(X_{ik} - \bar{X}(k))$$

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

la matriz de varianza y covarianza esta dada por

$$S = E[(X - \mu)(X - \mu)^T]$$

donde :

$$\mu = E[X] = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_p] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

$$S = E \left[\begin{bmatrix} (X_1 - \mu_1) \\ (X_2 - \mu_2) \\ \vdots \\ (X_p - \mu_p) \end{bmatrix} \begin{bmatrix} (X_1 - \mu_1) & (X_2 - \mu_2) & \cdots & (X_p - \mu_p) \end{bmatrix} \right]$$

$$S = \begin{bmatrix} E(X_1 - \mu_1)^2 & E(X_1 - \mu_1)(X_2 - \mu_2) & \cdots & E(X_1 - \mu_1)(X_p - \mu_p) \\ E(X_2 - \mu_2)(X_1 - \mu_1) & E(X_2 - \mu_2)^2 & \cdots & E(X_2 - \mu_2)(X_p - \mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_p - \mu_p)(X_1 - \mu_1) & E(X_p - \mu_p)(X_2 - \mu_2) & \cdots & E(X_p - \mu_p)^2 \end{bmatrix}$$

$$S = \text{Cov}(x) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{bmatrix}$$

donde $\sigma_{ii} = \sigma_i^2 = Var(X_i)$ $i=1,2,\dots,p$

2.2.4 MATRIZ DE CORRELACIÓN

A partir de la estandarización de los elementos de la matriz S da lugar a las correlaciones, y es posible calcular los elementos de la matriz R, siendo ambas de igual dimensión y cuyos elementos sean los coeficientes de correlación entre la j-ésima y la k-ésima variable, Y esta definido de la siguiente forma:

$$r(jk) = \frac{S(jk)}{\sqrt{S(ij)S(kk)}} = \frac{S(jk)}{S(j)S(k)}$$

O también la podemos representar de la siguiente manera

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}$$

podemos probar siempre que

$$-1 \leq \rho_{ij} \leq 1$$

una vez obtenida la matriz de varianza y covarianza podemos calcular la matriz de correlación ρ , la misma que se define de la siguiente forma:

$$\rho = \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{11}}} & \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \dots & \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} \\ \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \frac{\sigma_{22}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{22}}} & \dots & \frac{\sigma_{2p}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{pp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} & \frac{\sigma_{2p}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{pp}}} & \dots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{pp}}} \end{bmatrix}$$

También se puede determinar mediante arreglos en una matriz de correlación muestral cuya diagonal principal estará formada por números uno y será simétrica como la matriz de covarianza por ser $r(jk)=r(kj)$:

$$R = \begin{bmatrix} 1 & \dots & \dots & r(1p) \\ \vdots & 1 & \vdots & \vdots \\ \vdots & \vdots & 1 & \vdots \\ r(p1) & \vdots & r(pk) & 1 \end{bmatrix}$$

2.2.5 EL METODO DE ANALISIS POR COMPONENTES PRINCIPALES

El análisis de componentes principales es una técnica multivariada de interdependencia, también es una variación del análisis de factores, permite explicar la matriz de varianzas y covarianzas de un

conjunto de variables, a través de las combinaciones lineales de variables observables, en la que se estudian p variables de interés, que constituyen un vector aleatorio $X = X_1, X_2, \dots, X_p$, posiblemente normal multivariado ($X \sim N(\mu, \Sigma)$) o tal vez en el que estas p variables observables (componentes), generan k (componentes principales) variables latentes $k < p$, que se pretende que contengan tanta información como sea posible para así lograr reducir los datos.

Los objetivos son:

- Reducir los datos del problema que se está estudiando.
- Crear nuevas variables que expresen la misma información contenida en el conjunto original.
- Eliminar ciertas variables originales que prestan poca información
- Descubrir relaciones entre variables que muchas veces no son consideradas para el análisis.

Este análisis se aplica cuando se dispone de un conjunto de datos multivariados y no se puede postular, sobre la base de conocimientos previos del universo en estudio.

Además, el análisis de componentes principales se aplicará cuando se desee conocer la relación entre los elementos de una población, y se crea que en dicha relación influye de manera desconocida un conjunto de variables.

Sea X un vector aleatorio en R^p variado siendo $X^T = (X_1, X_2, \dots, X_p)$ que tiene la matriz de covarianza Σ , con valores propios $\lambda_1 \geq \lambda_2 \geq \dots \lambda_p \geq 0$ obtenidos a partir de la ecuación:

$$\det(\Sigma - \lambda I) = 0$$

definamos p variables no observables Y_1, Y_2, \dots, Y_p como una combinación lineal de X_1, X_2, \dots, X_p así.

$$\begin{aligned} Y_1 &= d_1^T X = d_{11}X_1 + d_{12}X_2 + \dots + d_{1p}X_p \\ Y_2 &= d_2^T X = d_{21}X_1 + d_{22}X_2 + \dots + d_{2p}X_p \\ &\vdots \\ Y_p &= d_p^T X = d_{p1}X_1 + d_{p2}X_2 + \dots + d_{pp}X_p \end{aligned}$$

Obteniendo así la varianza y la covarianza de cada una de las combinaciones lineales esta dada por:

$$\text{Var}(Y_i) = \text{Var}(d_i^T X) = d_i^T \Sigma d_i \quad i=1,2,\dots,p$$

$$\text{Cov}(Y_i, Y_k) = d_i^T \Sigma d_k \quad i=1,2,\dots,p$$

Las componentes principales son combinaciones lineales no correlacionadas, cuyas varianzas son muy grandes. La primera componente principal es la máxima combinación lineal con la máxima varianza.

$d_1^T X$ donde $\text{Var}(d_1^T X)$ está sujeta a $d_1^T d_1 = 1$

La segunda componente principal es la máxima combinación lineal de $d_2^T X$ donde.

$\text{var}(d_2^T X)$ está sujeta a $a_2^T a_2 = 1$ y $\text{Cov}(d_1^T X, d_2^T X) = 0$

y en la i -ésima componente será la máxima combinación lineal de $d_i^T X$ donde, $\text{Var}(d_i^T X)$ está sujeta a.

$$d_i^T d_i = 1 \quad \text{y} \quad \text{Cov}(d_i^T X, d_k^T X) = 0 \quad \text{para} \quad \frac{\lambda_i}{(\lambda_1 + \lambda_2 + \dots + \lambda_p)}$$

Dada la matriz de covarianza asociada al vector aleatorio

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

De los cuales podemos obtener los pares de los valores característicos denotados como

$$(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p) \text{ donde } \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \text{ lo cual demuestra}$$

que la i -ésima componente principal esta denotada por.

$$Y_i = e_i^T X = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p; \quad i=1,2,\dots,p$$

Donde la varianza de cada componente es igual al correspondiente valor propio, es decir que se puede probar que:

$$Var(Y_i) = e_i^T \Sigma e_i = \lambda_i; \quad i=1,2,\dots,p$$

Además las componentes principales son variables no correlacionadas

$$Cov(Y_i, Y_k) = e_i^T \Sigma e_k = 0 \quad i \neq k$$

Cuando $a = e_1^T \quad e_1^T e_1 = 1$ ya que los vectores característicos son normalizados tenemos.

$$\max_{a \neq 0} [d^T \Sigma d / d^T d] = \lambda_1 = e_1^T \Sigma e_1 / e_1^T e_1 = \text{var}(Y_1)$$

Generalizando para los demás vectores característicos se representa como:

$$\max_{a / e_1, e_2, \dots, e_k} [a^T \Sigma a / a^T a] = \lambda_{k+1} ; \quad k=1, 2, \dots, p-1$$

en el caso particular $a = e_{k+1}$, con $e_{k+1}^T e_i = 0$

para $i=1, 2, \dots, k$ y $k=1, 2, \dots, p-1$

tenemos:

$$e_{k+1}^T \Sigma e_{k+1} / e_{k+1}^T e_{k+1} = e_{k+1}^T \Sigma e_{k+1} = \text{var}(Y_{k+1}) = \lambda_{k+1}$$

esto es $e_i^T e_k = 0$ para $i \neq k$ obteniéndose $\text{Cov}(Y_i, Y_k) = 0$

como podemos observar los valores propios de Σ son ortogonales y todos los valores propios $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_p$ son distintos. En el caso de que los valores propios no fueran distintos, los vectores propios correspondientes para cada uno de los valores propios comunes pueden ser ortogonales.

Aquí para cualquier dos eigenvectores e_i y e_k , $e_i^t e_k = 0$ para $i \neq k$

$\Sigma e_k = \lambda_k e_k$ multiplicando por e^t .

$$\text{Cov}(Y_i, Y_k) = e_i^t \Sigma e_k = e_i^t \lambda_k e_k = \lambda_k e_i^t e_k = 0 \quad \text{para } i \neq k$$

también se puede probar que :

$$\sigma_{11} + \sigma_{22} + \dots + = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_p)$$

$$= \lambda_1 + \lambda_2 + \dots + \lambda_p$$

$$= \text{Var}(Y_1) + \text{Var}(Y_2) + \dots + \text{Var}(Y_p)$$

La proporción de la varianza total explicada por la i -ésima componente principal esta dada por:

$$\frac{\lambda_i}{(\lambda_1 + \lambda_2 + \dots + \lambda_p)} \quad i=1,2,\dots$$

El porcentaje de varianza que se desee explicar, determinará el número de componentes principales que se escogerá para el análisis del estudio que se este realizando.

Si una gran parte de la población puede ser explicada por una, dos o tres componentes, entonces, esta puede reemplazar las p variables originales sin mucha perdida de información.

CAPITULO III

ANÁLISIS ESTADÍSTICO

INTRODUCCIÓN

La producción de caña de azúcar en el mercado nacional, al igual que su cultivo esta influenciado por variables agrícolas.

Uno de los ingenios como lo es Valdez quien nos a proporcionado la información de los datos como superficie sembrada, cosecha, producción, roza de cantero, herbicida, rastrillos, escarificado, riego, sacada de paja cauca, aporque, fertilización.

3.1 Variables a utilizar

3.1.1 Superficie Sembrada

La cuales son representadas mediante tierras ocupadas por semillas o plantas cuyo objetivo es producir, la producción agrícola de caña de azúcar en los periodos de referencia, la cual se mide en hectáreas (Has).

3.1.2 Superficie Cosechada

Son aquellas tierras que luego de ser sembradas han sido cosechadas generando producción agrícola dentro de los periodos de referencia, esta área puede ser menor o igual al área sembrada. Se mide en hectáreas(Has).

3.1.3 Producción

La producción agrícola es la cantidad total del producto primario en este caso la caña de azúcar obtenida del cultivo en el periodo de referencia. Se mide en toneladas métricas (Tons).

3.1.4 Roza de Canteros

Es la labor de controlar el crecimiento de la maleza, en este caso monte o paja, utilizando herramientas manuales. Se mide en hectáreas (Has)

3.1.5 Herbicidas

Consiste en la aplicación de productos químicos para el control de la maleza. Se mide en hectáreas (Has).

3.1.6. Rastrillo

Recolección de la basura o de la hojarasca que queda luego de la cosecha, para dejar limpia el área y permitir que se pueda desarrollar la planta. Estos restos son apilados o amontonados para ser luego quemados. Se mide en hectáreas (Has).

3.1.7 Escarificado

Es la labor de aflojar o ablandar el terreno que ha sido compactado por el tráfico pesado de vehículos y maquinarias utilizados en la cosecha, con el propósito de que el agua de riego penetre y se conserve por más tiempo la humedad y para la labor de fertilización. Se mide en hectáreas (Has).

3.1. 8 Riego

GRAFICO XI: RIEGO



Son las labores relacionadas con la provisión de agua necesaria para el desarrollo de los cultivos (riego) y la evacuación del exceso de agua en la época invernal para evitar inundaciones y daños en los cultivos. Esta se mide en hectáreas (Has).

3.1.9 Sacada de paja cauca

Llamado también Control manual de malezas, este consiste en la labor de controlar el crecimiento de la maleza (especialmente sacando la paja cauca), utilizando herramientas manuales. Esta se mide en hectáreas (Has).

3.1.10 Aporque

También llamado cultivo mecánico, es la labor de acumular tierra alrededor de la planta para que se pueda nutrir mejor, que se mantenga erecta y pueda desarrollarse normalmente. Esta se mide en hectáreas(Has).

3.1.11 Fertilización

Esta labor consiste en la aplicación de fertilizantes para mejorar la calidad del suelo y lograr mejores rendimientos de la caña. Esta se mide en hectáreas(Has).

3.2 ESTADÍSTICA DESCRIPTIVA DE LAS VARIABLES DE ESTUDIO

3.2.1 Superficie Sembrada

La superficie sembrada de caña de azúcar en el periodo del año 1980 hasta el año 2020 no ha variado significativamente según los registros analizados.

A continuación tenemos la estadística descriptiva de esta variable.

La superficie sembrada mínima fue en el año 1989 y la máxima superficie sembrada tuvo lugar en el año 2002, esta variable tiene una media de 10319, y sus datos se encuentran dispersos en relación a la media 821, has.

TABLA VIII: Estadística descriptiva: Variable Superficie Sembrada

N de Casos	14
Mínimo(1989)	8221,77
Máximo(2002)	11114,23
Rango	2892,46
Mediana	10490
Media	10319,27
Moda	10380
Nivel de confianza(95,0%)	474,2886127
Error estándar	219,5478753
Desviación estándar	821,4729295
Varianza	674817,7739
sesgo	-1,72

Fuente: Ingenio Valdez

En la figura 3.1 podemos ilustrar el histograma de frecuencias de la variable el cual se encuentra detallado en la tabla III. La clase modal se encuentra en el intervalo **[10080-10550)**, cuyo intervalo es el que contiene el mayor número de observaciones.

**TABLA IX : TABLA DE FRECUENCIAS DE SUPEFICIE
SEMBRADA**

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada Relativa
[8200-8670)	8435	1	0,071	0,071
[8670-9140)	8905	1	0,071	0,143
[9140-9610)	9375	0	0,000	0,143
[9610-10080)	9845	1	0,071	0,214
[10080-10550)	10315	7	0,500	0,714
[10550-11020)	10785	2	0,143	0,857
[11020-11490)	11255	2	0,143	1,000

El **intervalo de clase** está conformado por los valores tomados en este caso de la superficie sembrada donde estos valores se les a determinado un valor de separación para poder tomar su frecuencia absoluta de una manera adecuada, el valor es 470.

Marca de clase es el valor resultante de la suma del intervalo dividido para dos como tenemos en la tabla de frecuencia $(8200+8670)/2$.

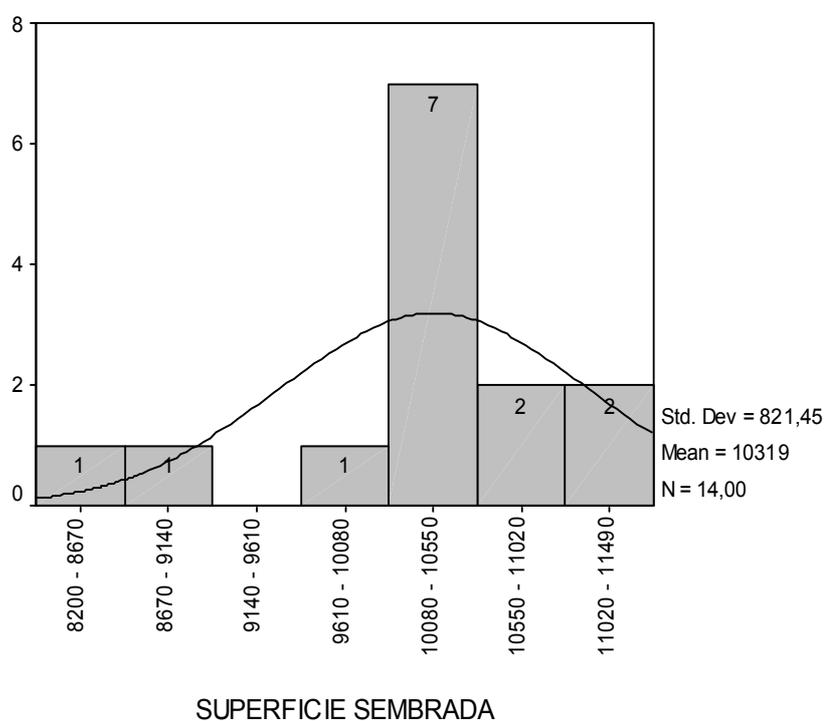
Frecuencia absoluta es el número de observaciones que se encuentran en cada intervalo de clase.

Frecuencia relativa es el valor de la frecuencia absoluta dividida para el número total de observaciones ejemplo $1/14$.

Frecuencia acumulada relativa es la suma acumulada de la frecuencia relativa ejemplo $0.071+0.071\dots$

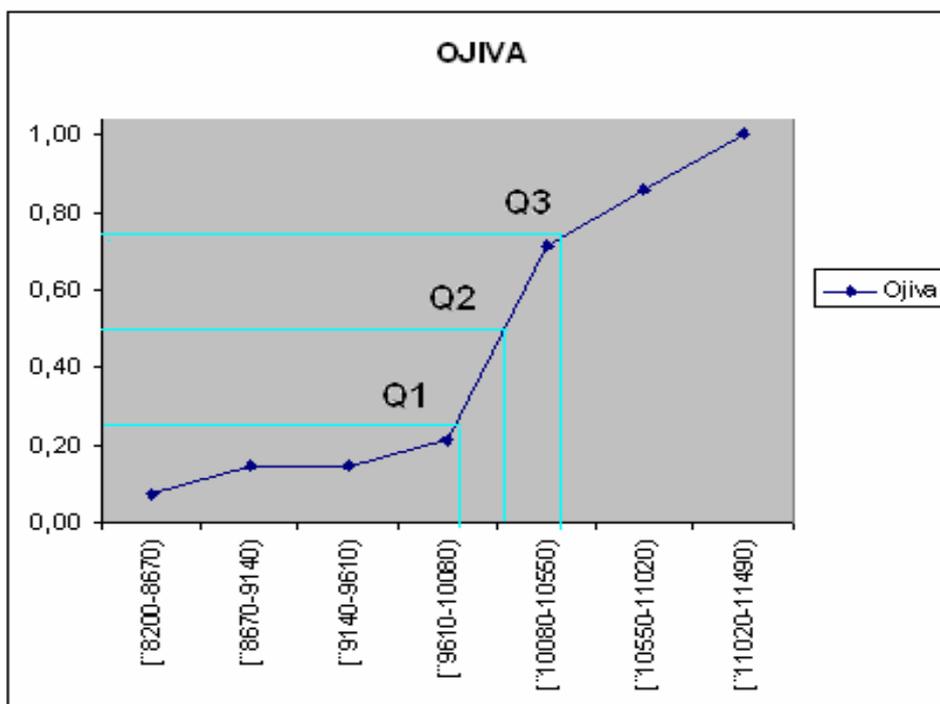
GRAFICO XII: HISTOGRAMA DE FRECUENCIAS DE SUPERFICIE

SEMBRADA



El Gráfico XII nos muestra que la superficie sembrada con mayor frecuencia está en el intervalo 10080-10550 durante los años 1989-2002.

GRAFICO XIII: OJIVA DE SUPERFICIE SEMBRADA



Los valores que nos representan Q1, Q2 y Q3 que dividen a los datos colocados por orden de magnitud en cuatro partes iguales son denominados cuartiles. Como podemos observar en la figura 3.2 el primer cuartil cuyo valor es 10378 cae en el quinto intervalo de clase, el segundo cuartil el cual es la mediana que es la mejor medida de posiciones en distribuciones sesgadas es 10489 y el tercer cuartil corresponde al valor 10500. Por lo tanto el alcance intercuartil que significa que tan lejos de la mediana tenemos que ir en cualquiera de

las dos direcciones antes de recorrer la mitad de los valores del conjunto de datos es $Q_3 - Q_1$ y en este caso es de 736,23 .

El 25% de la superficie sembrada es menor o igual que 10378 hás, el 50% de la superficie sembrada es menor o igual que 10489 y el 75% de la superficie sembrada es menor o igual que 10500 hás.

Tenemos el intervalo de clase modal que es $[10080-10550)$ donde se encuentra concentrado la mayor cantidad de datos.

3.2.2 SUPERFICIE COSECHADA

Los datos de esta variable corresponden al período de los años 1989 hasta el 2002. Esta variable está en miles de hectáreas.

La cantidad mínima cosechada fue en el año 1997, que si recordamos coincide con el año que se vio afectada la agricultura por el Fenómeno del Niño, la mejor cosecha fue realizada en el año 1999.

TABLA X: Estadística descriptiva: Variable Superficie cosechada

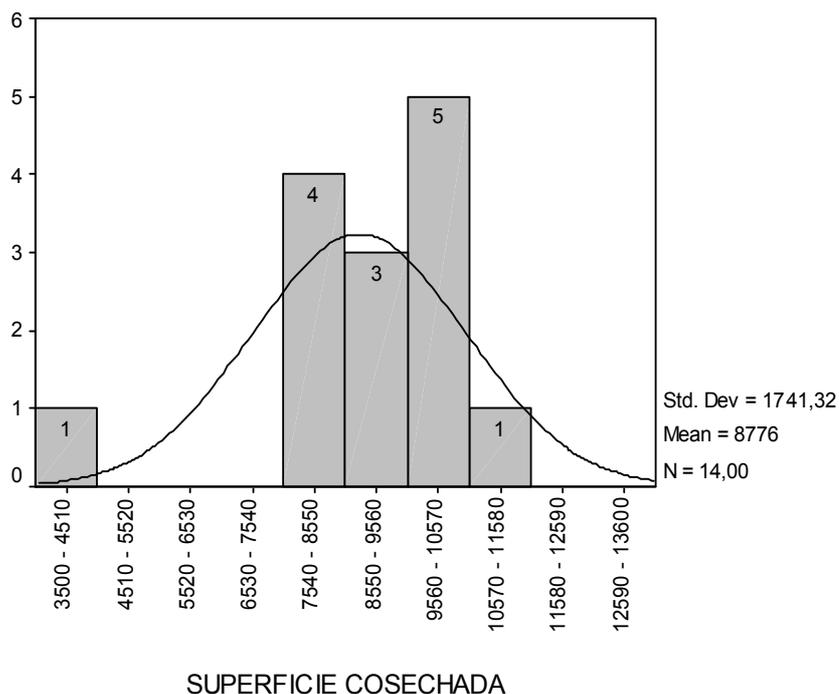
N de Casos	14
Mínimo(1997)	3548,77
Máximo(1999)	9561044
Rango	9557495,23
Mediana	9119,11
Media	691024,87
Nivel de confianza(95,0%)	1474039,384
Error estándar	682309,32
Desviación estándar	2552967,71
Varianza	6,51764E+12
sesgo	3,741654391

Fuente: Ingenio Valdez

TABLA XI TABLA DE FRECUENCIAS DE SUPERFICIE COSECHADA

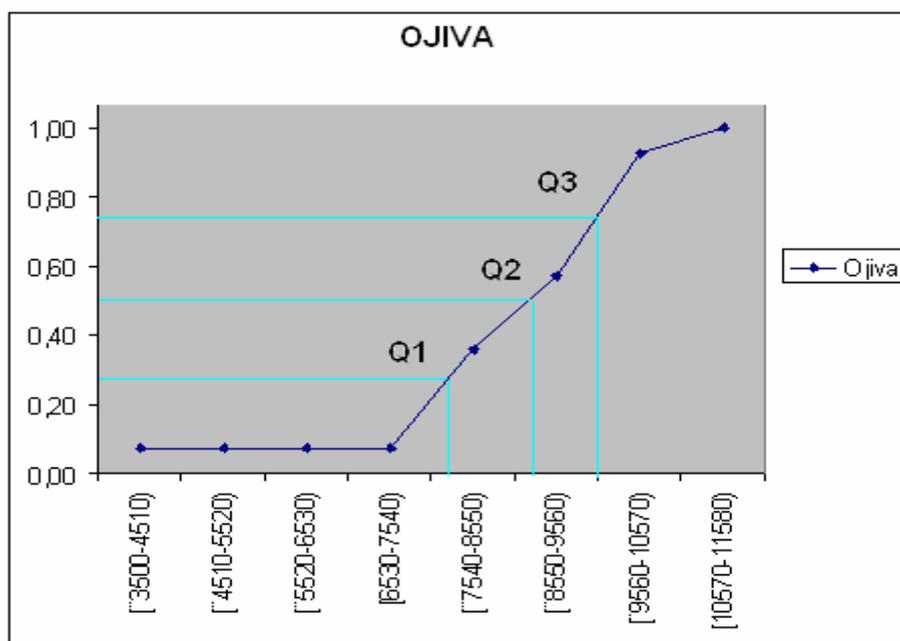
Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[3500-4510)	4005	1	0,071	0,071
[4510-5520)	5015	0	0,000	0,071
[5520-6530)	6025	0	0,000	0,071
[6530-7540)	7035	0	0,000	0,071
[7540-8550)	8045	4	0,286	0,357
[8550-9560)	9055	3	0,214	0,571
[9560-10570)	10065	5	0,357	0,929
[10570-11580)	11075	1	0,071	1,000

**GRAFICO XIV: HISTOGRAMA DE FRECUENCIAS DE
SUPERFICIE COSECHADA**



El grafico mostrado en la Grafico XIV nos muestra que la cantidad que se cosecha con mayor frecuencia dentro de los años de estudio en este caso 1989-2002 esta en los intervalos 9560-10570 , siendo de igual manera la menor producción en el intervalo 3500-4510 debido que justamente en este año fue afectada por el fenómeno del niño.

GRAFICO XV: OJIVA DE SUPERFICIE COSECHADA



En el Grafico XV se muestra el polígono de frecuencia acumulado u ojiva. El primer cuartil corresponde al valor 8080,16 el segundo cuartil corresponde 9119,11 que a su vez es la mediana y el tercer cuartil es 9769,17. El intervalo correspondiente a la clase modal es [9560-10570) como se ilustran en los gráficos en el eje de las X se encuentra la marca de clase para cada uno de los intervalos.

3.2.3 PRODUCCIÓN

Los datos de esta variable corresponden desde el año 1989 hasta el año 2002 en forma anual.

La mínima producción corresponde al año de 1997 alcanzando la máxima en el año 1998.

**TABLA XII: ESTADÍSTICA DESCRIPTIVA: VARIABLE
PRODUCCIÓN**

N de Casos	14
Mínimo(1997)	405631,13
Máximo(1998)	865019,75
Rango	459388,62
Mediana	730481
Media	709029,8507
Nivel de confianza(95,0%)	65728,51084
Error estándar	30424,67929
Desviación estándar	113838,726
Varianza	1295925537
sesgo	-1,484137227

Fuente: Ingenio Valdez

TABLA XIII: TABLA DE FRECUENCIAS DE LA PRODUCCIÓN

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[405000-475100)	440050	1	0,0714	0,0714
[475100-545200)	510150	0	0,0000	0,0714
[545200-615300)	580250	1	0,0714	0,1428
[615300-685400)	650350	2	0,1428	0,2857
[685400-755500)	720450	6	0,4285	0,7142
[755500-825600)	790550	2	0,1428	0,8571
[825600-895700)	860650	2	0,1428	1,0000

GRAFICO XVI: HISTOGRAMA DE FRECUENCIAS DE LA PRODUCCIÓN

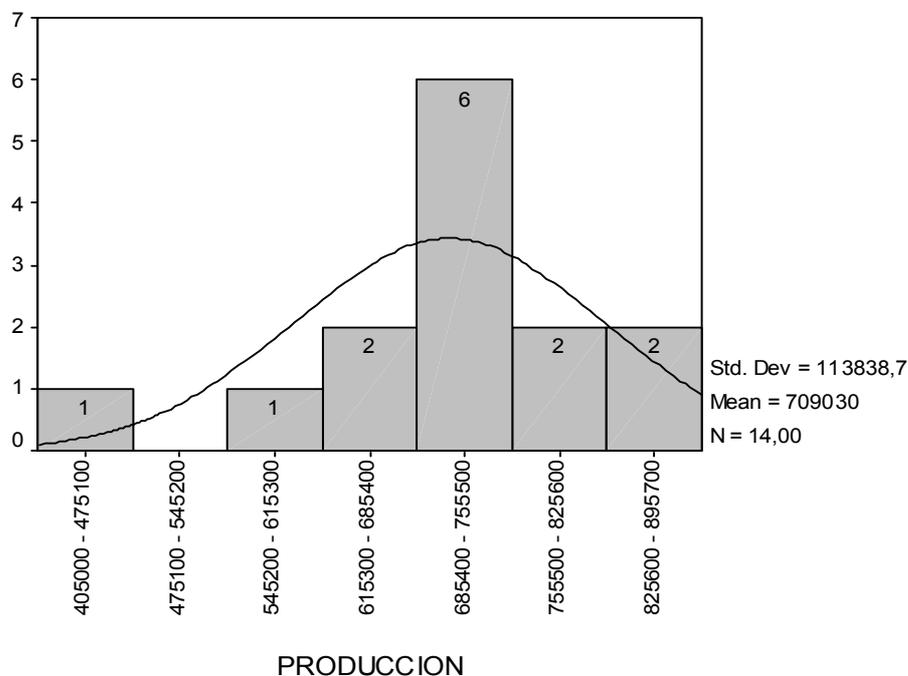
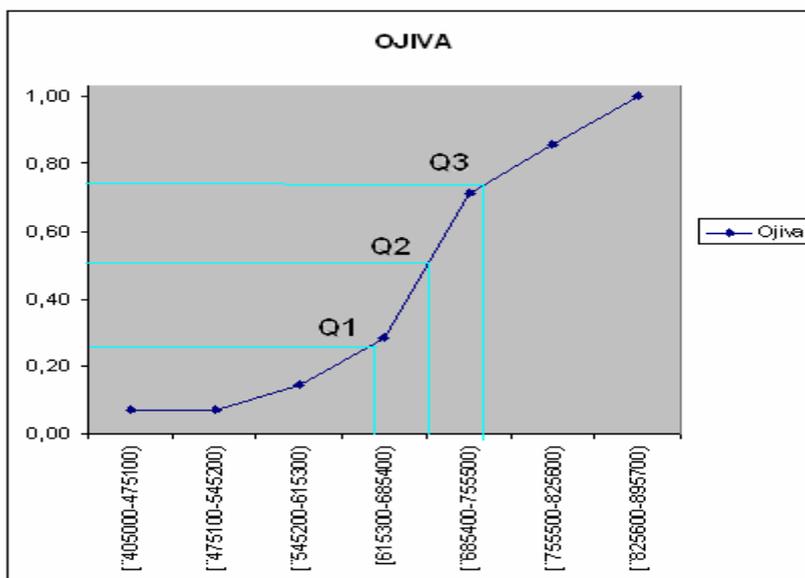


GRAFICO XVII: OJIVA DE LA PRODUCCIÓN



En el Gráfico XVII se muestra el gráfico que muestra las frecuencias acumuladas menores que el límite real superior de la clase.

Para una mejor ilustración tenemos un ejemplo la frecuencia acumulada relativa menor que 755500 es 0.69 lo cual interpreta el 69% de la producción en el periodo de los años 1989 -2002 tiene menos de 755.500 Tm.

3.2.4 ROZA DE CANTEROS

Los datos de esta variable corresponden desde el año 1989 hasta el año 2002 en forma anual. Esta variable está en miles de hectáreas, y además ilustraremos mediante un indicador el cual nos mostrara el número de veces de la actividad realizada.

**TABLA XIV: ESTADÍSTICA DESCRIPTIVA: VARIABLE ROZA
DE CANTERO**

N de Casos	14
Mínimo(1997)	345,2
Máximo(1998)	22424,18
Rango	22078,98
Mediana	20583,7
Media	18352,105
Nivel de confianza(95,0%)	3435,36949
Error estándar	1590,177742
Desviación estándar	5949,900295
Varianza	35401313,52
sesgo	-2,438414297

Fuente: Ingenio Valdez

GRAFICO XVIII: HISTOGRAMA DE FRECUENCIAS DE ROZA DE CANTERO

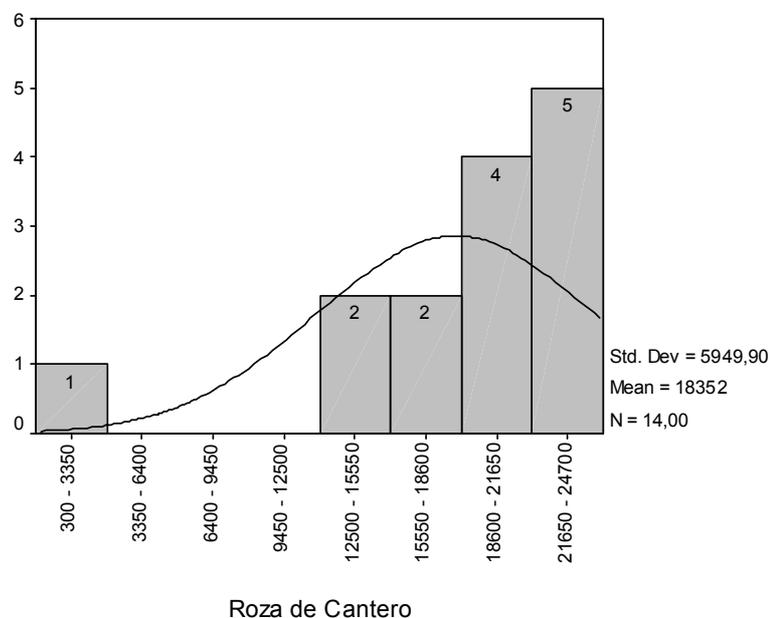
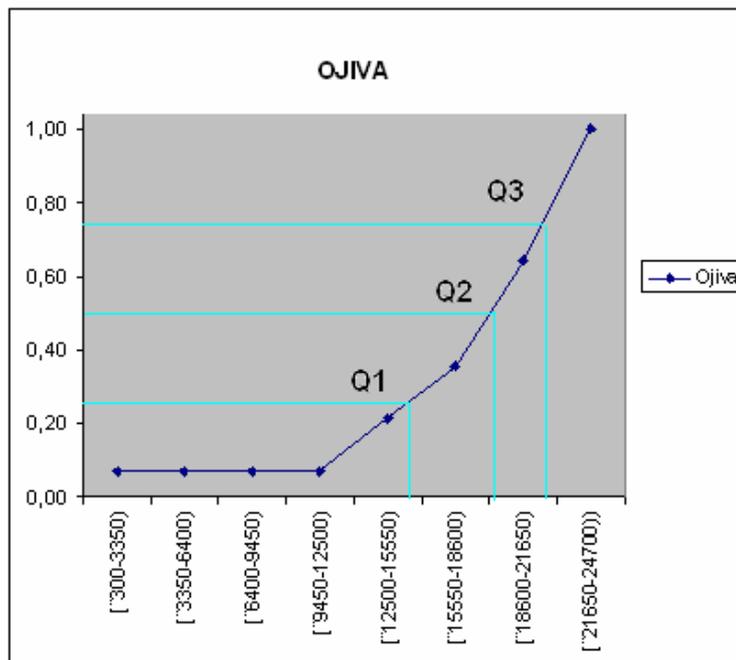


TABLA XV: TABLA DE FRECUENCIAS DE ROZA DE CANTERO

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[300-3350)	1825	1	0,071	0,071
[3350-6400)	4875	0	0,000	0,071
[6400-9450)	7925	0	0,000	0,071
[9450-12500)	10975	0	0,000	0,071
[12500-15550)	14025	2	0,143	0,214
[15550-18600)	17075	2	0,143	0,357
[18600-21650)	20125	4	0,286	0,643
[21650-24700))	23175	5	0,357	1,000

GRAFICO XIX: OJIVA DE ROZA DE CANTERO



La frecuencia acumulada de la variable roza de cantero muestra que el 25% de los datos Q1 es menor e igual 16.726 hás donde Q2 es 20584 que es igual a su media y Q3 es 21806 siendo el valor que representa el 75% de la roza de cantero cuyo valor es 21806.

Esta variable también la podemos analizar mediante un **indicador** el cual nos determina las veces en la que las hectárea sembradas fueron rozada.

Los datos de esta variable corresponden desde el año 1989 hasta el año 2002 en forma anual.

3.2.5 INDICADOR DE ROZA DE CANTERO

TABLA XVI: ESTADÍSTICA DESCRIPTIVA: VARIABLE ROZA DE CANTERO MEDIANTE UN INDICADOR DE VECES

N de Casos	14
Mínimo(2000)	0,03
Máximo(1992)	2,12
Rango	2,09
Mediana	1,995
Moda	1,98
Media	1,789285714
Nivel de confianza(95,0%)	0,324098474
Error estándar	0,150020014
Desviación estándar	0,561323495
Varianza	0,315084066
sesgo	-2,80959628

Fuente: Ingenio Valdez

GRAFICO XX: HISTOGRAMA DE FRECUENCIA DE ROZA DE CANTERO MEDIANTE UN INDICADOR DE VECES

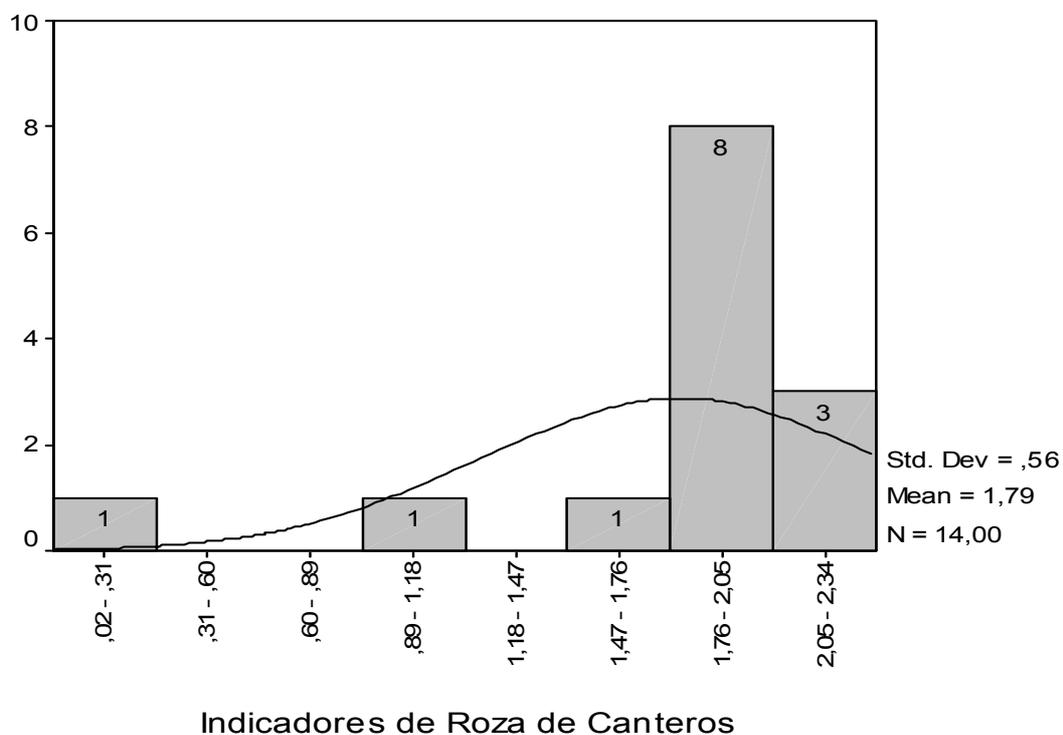
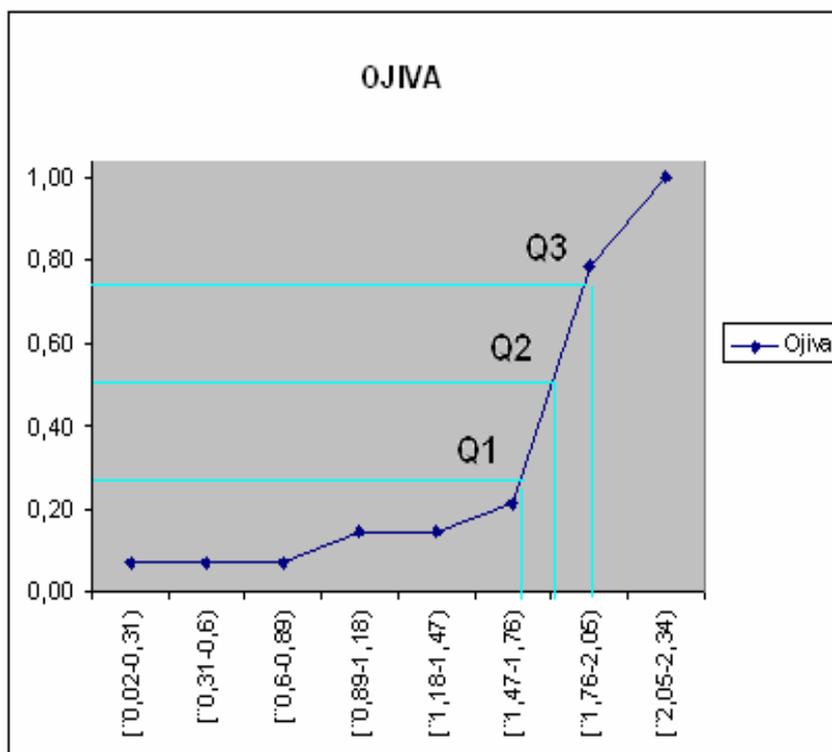


TABLA XVII: TABLA DE FRECUENCIA DE ROZA DE CANTERO MEDIANTE UN INDICADOR DE VECES

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[0,02-0,31)	0,165	1	0,07142	0,07142
[0,31-0,6)	0,455	0	0,00000	0,07142
[0,6-0,89)	0,745	0	0,00000	0,07142
[0,89-1,18)	1,035	1	0,07142	0,14285
[1,18-1,47)	1,325	0	0,00000	0,14285
[1,47-1,76)	1,615	1	0,07142	0,21428
[1,76-2,05)	1,905	8	0,57142	0,78571
[2,05-2,34)	2,195	3	0,21428	1,00000

GRAFICO XXI: OJIVA ROZA DE CANTERO MEDIANTE UN INDICADOR DE VECES



La frecuencia acumulada de la variable rendimiento muestra que el 25% de los datos Q1 es menor o igual que 1.83 casi 2 veces, el 50% Q2 es menor o igual que 1.99 aproximadamente 2 veces y el 75% Q3 es menor igual que 2.02 que es exactamente 2 veces.

3.2.6 HERBICIDAS

**TABLA XVIII: ESTADÍSTICA DESCRIPTIVA: VARIABLE
HERBICIDAS**

N de Casos	14
Mínimo(2001)	3531
Máximo(1996)	17475
Rango	13944
Mediana	13510,97
Media	12377,81714
Nivel de confianza(95,0%)	2266,555653
Error estándar	1049,15246
Desviación estándar	3925,569051
Varianza	15410092,38
sesgo	-0,96707303

Fuente: Ingenio Valdez

**GRAFICO XXII: HISTOGRAMA DE FRECUENCIA DE
HERBICIDAS**

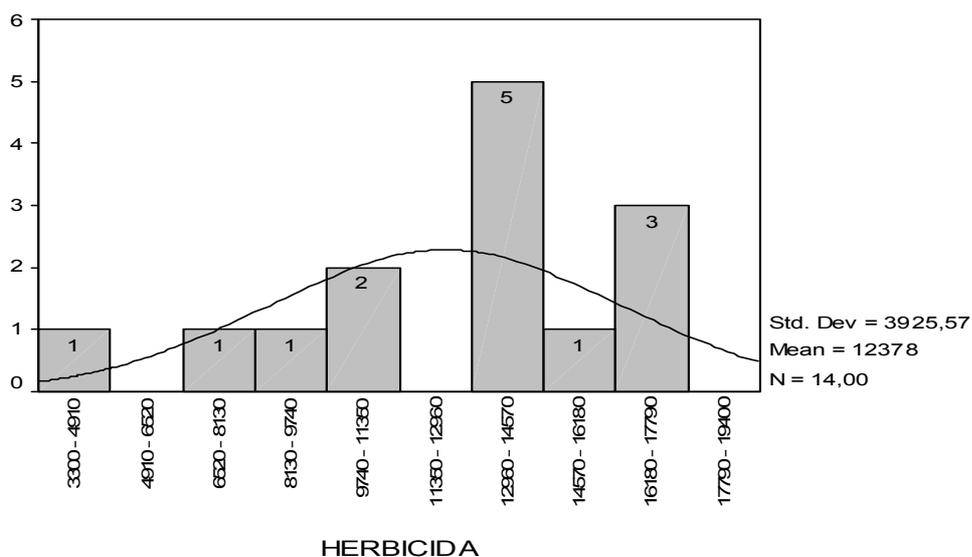
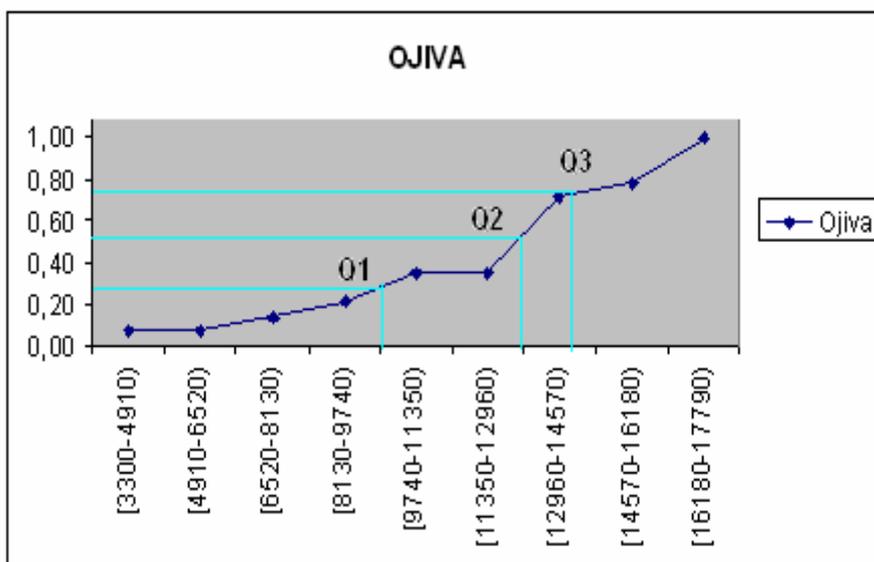


TABLA XIX: TABLA DE FRECUENCIA DE HERBICIDAS

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[3300-4910)	4105	1	0,0714	0,0714
[4910-6520)	5715	0	0,0000	0,0714
[6520-8130)	7325	1	0,0714	0,1428
[8130-9740)	8935	1	0,0714	0,2142
[9740-11350)	10545	2	0,1428	0,3571
[11350-12960)	12155	0	0,0000	0,3571
[12960-14570)	13765	5	0,3571	0,7142
[14570-16180)	15375	1	0,0714	0,7857
[16180-17790)	16985	3	0,2142	1,0000

GRAFICO XXIII: OJIVA DE HERBICIDAS



El primer cuartil es 10147, el segundo cuartil o mediana es 13510,97, y el tercer cuartil es de 14419. El rango semicuartílico es $(Q3-Q1)/2$, por lo cual en este caso es de 2136.

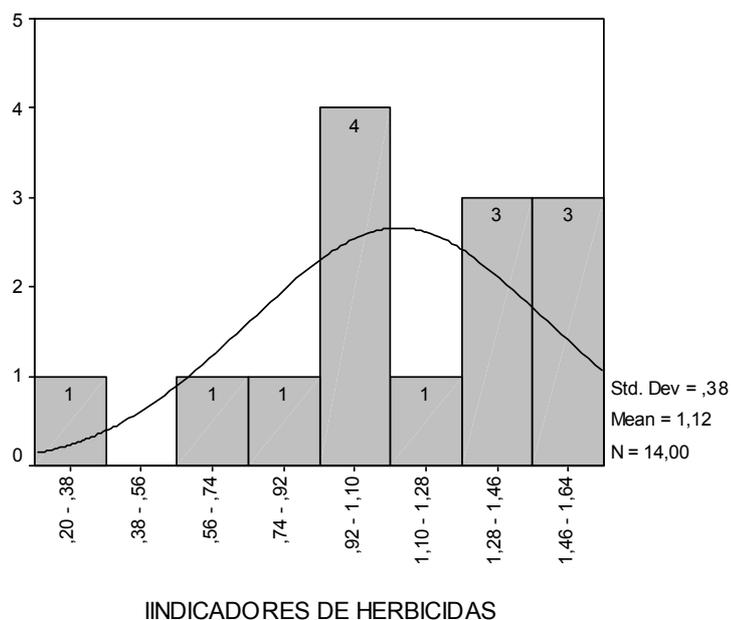
3.2.7 INDICADOR DE HERBICIDA

**TABLA XX: ESTADÍSTICA DESCRIPTIVA: VARIABLE
HERBICIDAS MEDIANTE UN INDICADOR DE VECES**

II de Casos	14
Mínimo(2000)	0,24
Máximo(1992)	1,57
Rango	1,33
Mediana	1,115
Moda	1,56
Media	1,120714286
Nivel de confianza(95,0%)	0,217768438
Error estándar	0,100801537
Desviación estándar	0,377164814
Varianza	0,142253297
sesgo	-0,875042868

Fuente: Ingenio Valdez

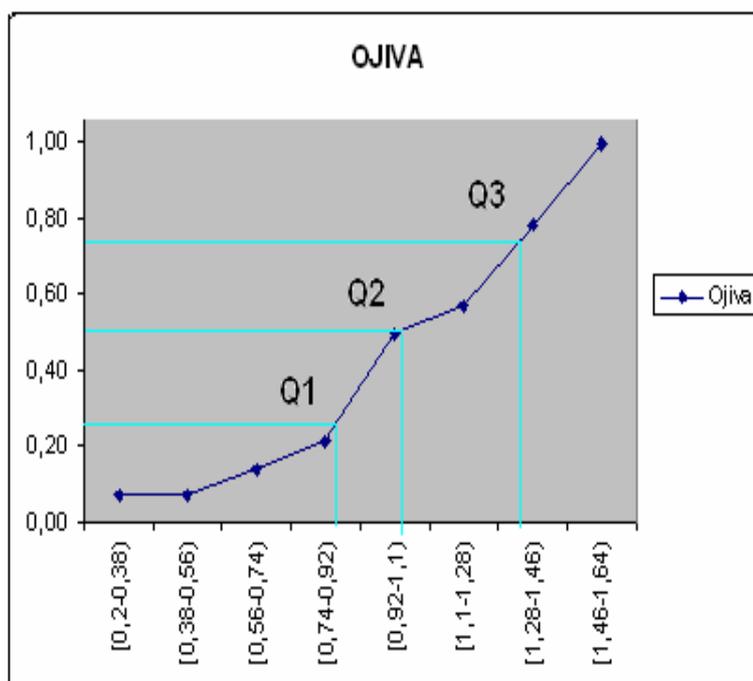
**GRAFICO XXIV: HISTOGRAMA DE FRECUENCIA DE
HERBICIDAS MEDIANTE UN INDICADOR DE VECES**



**TABLA XXI: TABLA DE FRECUENCIA DE HERBICIDAS
MEDIANTE UN INDICADOR DE VECES**

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[0,2-0,38)	0,29	1	0,0714286	0,0714286
[0,38-0,56)	0,47	0	0,0000000	0,0714286
[0,56-0,74)	0,65	1	0,0714286	0,1428571
[0,74-0,92)	0,83	1	0,0714286	0,2142857
[0,92-1,1)	1,01	4	0,2857143	0,5000000
[1,1-1,28)	1,19	1	0,0714286	0,5714286
[1,28-1,46)	1,37	3	0,2142857	0,7857143
[1,46-1,64)	1,55	3	0,2142857	1,0000000

**GRAFICO XXV: OJIVA DE HERBICIDAS MEDIANTE UN
INDICADOR DE VECES**



El primer cuartil indica el 25% de la aplicación de herbicidas es menor o igual que una vez, el segundo cuartil indica que el 50% de los datos es menor igual que 1.12 veces y el tercer cuartil indica que el 75% de los datos es menor o igual que 1.3 veces.

3.2.8 RASTRILLO

**TABLA XXII: ESTADÍSTICA DESCRIPTIVA: VARIABLE
RASTRILLO**

N de Casos	14
Mínimo(1997)	2769
Máximo(1996)	8770
Rango	6001
Mediana	7566,5
Media	7113,037857
Nivel de confianza(95,0%)	942,5791561
Error estándar	436,3048571
Desviación estándar	1632,503291
Varianza	2665066,996
sesgo	-1,421430061

Fuente: Ingenio Valdez

GRAFICO XXVI: HISTOGRAMA DE FRECUENCIA DE RASTRILLO

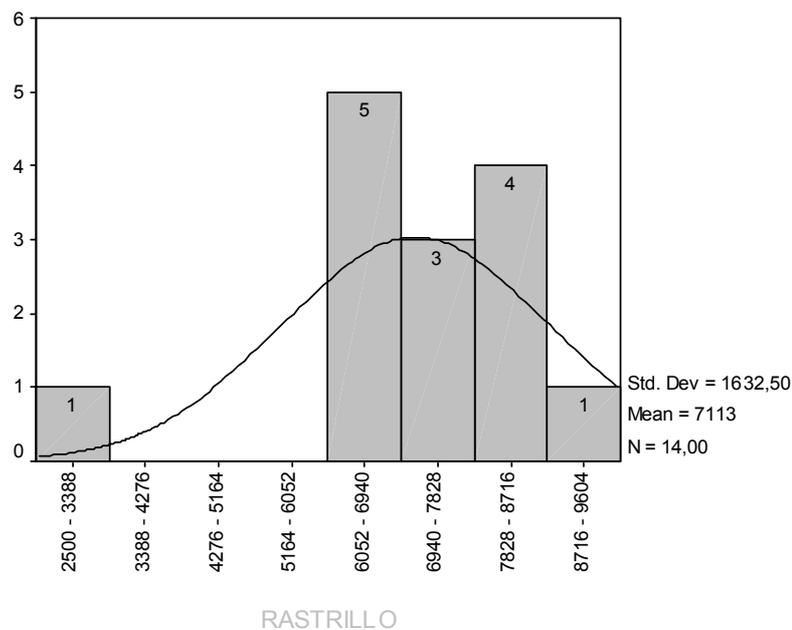
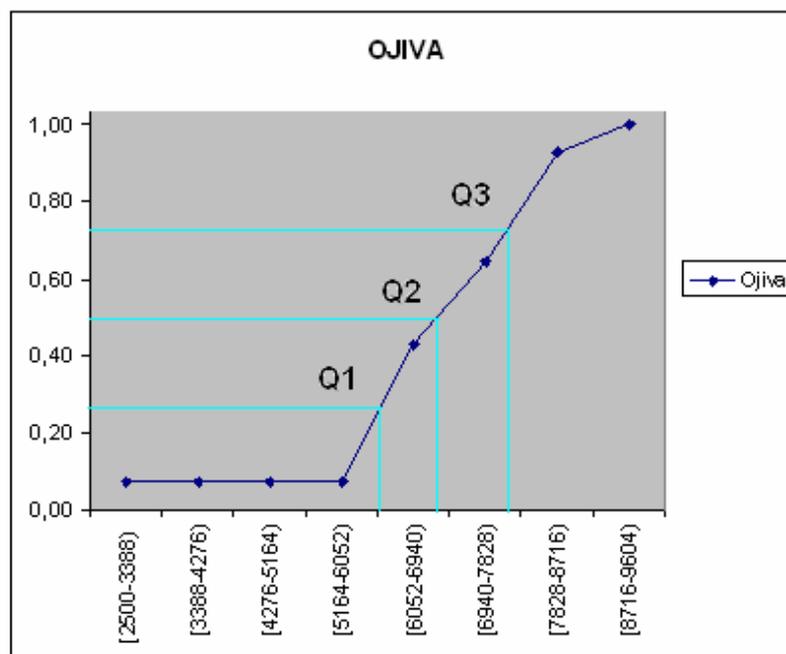


TABLA XXIII: TABLA DE FRECUENCIAS DE RASTRILLO

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[2500-3388)	2944	1	0,0714286	0,0714286
[3388-4276)	3832	0	0,0000000	0,0714286
[4276-5164)	4720	0	0,0000000	0,0714286
[5164-6052)	5608	0	0,0000000	0,0714286
[6052-6940)	6496	5	0,3571429	0,4285714
[6940-7828)	7384	3	0,2142857	0,6428571
[7828-8716)	8272	4	0,2857143	0,9285714
[8716-9604)	9160	1	0,0714286	1,0000000

GRAFICO XXVII: OJIVA DE RASTRILLO



En el Gráfico XXVII de la frecuencia acumulada de la figura 3.16 se puede ver que el primer cuartil es 6142, el segundo o mediana es 7566.5 y el tercer cuartil es de 8133.49.

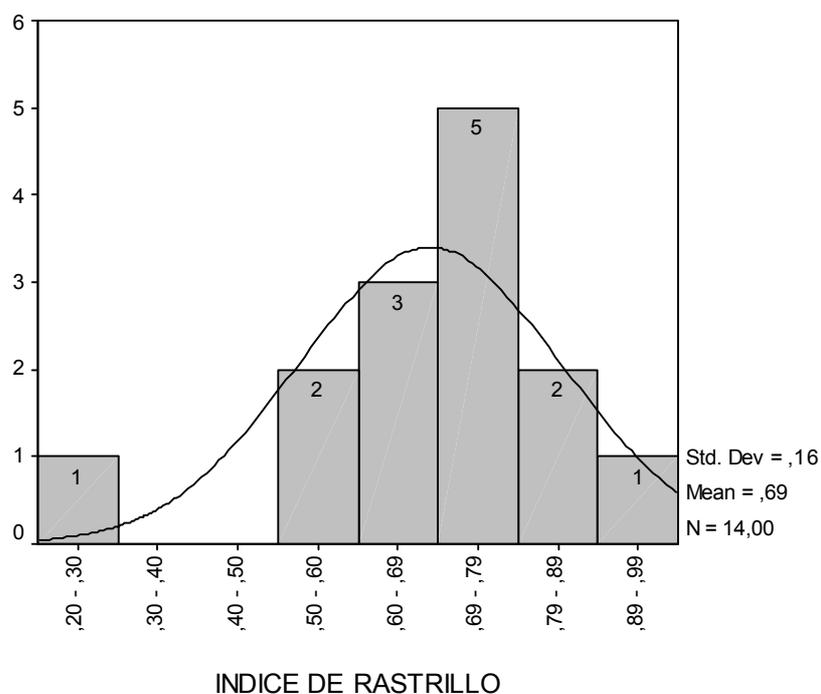
3.2.9 INDICADOR DE RASTRILLO

**TABLA XXIV: ESTADÍSTICA DESCRIPTIVA: VARIABLE
RASTRILLO MEDIANTE UN INDICADOR DE VECES**

N de Casos	14
Mínimo(1997)	0,25
Máximo(1996)	0,9
Rango	0,65
Mediana	0,725
Moda	0,59
Media	0,685
Nivel de confianza(95,0%)	0,093779409
Error estándar	0,043408993
Desviación estándar	0,162421579
Varianza	0,026380769
sesgo	-1,40603254

Fuente: Ingenio Valdez

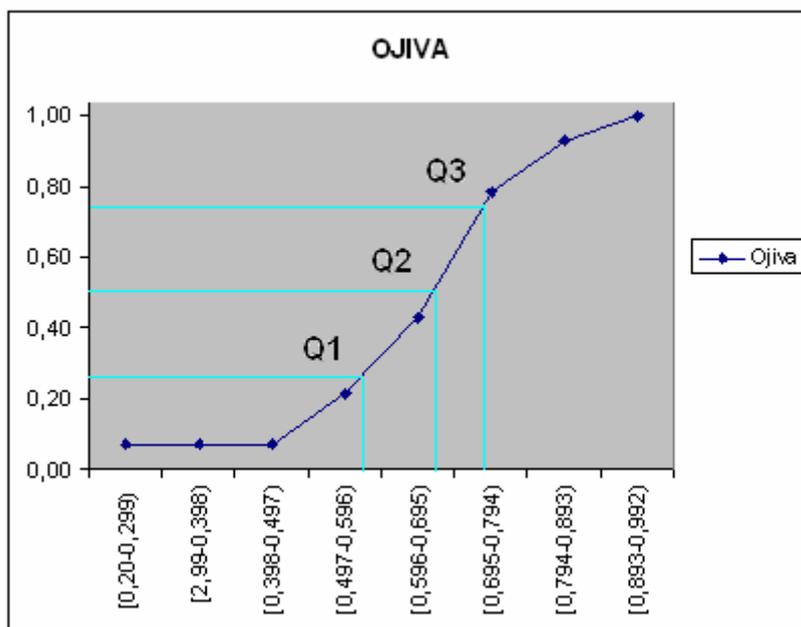
**GRAFICO XXVIII: HISTOGRAMA DE FRECUENCIAS RASTRILLO
MEDIANTE UN INDICADOR DE VECES**



**TABLA XXV: TABLA DE FRECUENCIA DE RASTRILLO
MEDIANTE UN INDICADOR DE VECES**

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[0,20-0,299)	0,2495	1	0,0714286	0,0714286
[2,99-0,398)	0,694	0	0,0000000	0,0714286
[0,398-0,497)	0,4475	0	0,0000000	0,0714286
[0,497-0,596)	0,5465	2	0,1428571	0,2142857
[0,596-0,695)	0,6455	3	0,2142857	0,4285714
[0,695-0,794)	0,7445	5	0,3571429	0,7857143
[0,794-0,893)	0,8435	2	0,1428571	0,9285714
[0,893-0,992)	0,9425	1	0,0714286	1,0000000

GRAFICO XXIX: OJIVA DE RASTRILLO MEDIANTE UN INDICADOR DE VECES



El primer cuartil es de 0.78 veces , el segundo cuartil es de 0.7 veces y el tercer cuartil es de 0.78 veces , lo que podemos dar nos cuenta que se irrealizado menos de 1 vez el rastriilo a las hectáreas que han sido cosechadas.

3.2.10 ESCLARIFICADO

**TABLA XXVI: ESTADÍSTICA DESCRIPTIVA: VARIABLE
ESCARIFICADO**

N de Casos	14
Mínimo(1997)	1448,65
Máximo(1996)	8607
Rango	7158,35
Mediana	7305,535
Media	6809,126429
Nivel de confianza(95,0%)	1027,605637
Error estándar	475,662259
Desviación estándar	1779,765205
Varianza	3167564,184
sesgo	-2,225116864

Fuente: Ingenio Valdez

**GRAFICO XXX: HISTOGRAMA DE FRECUENCIA DE
ESCARIFICADO**

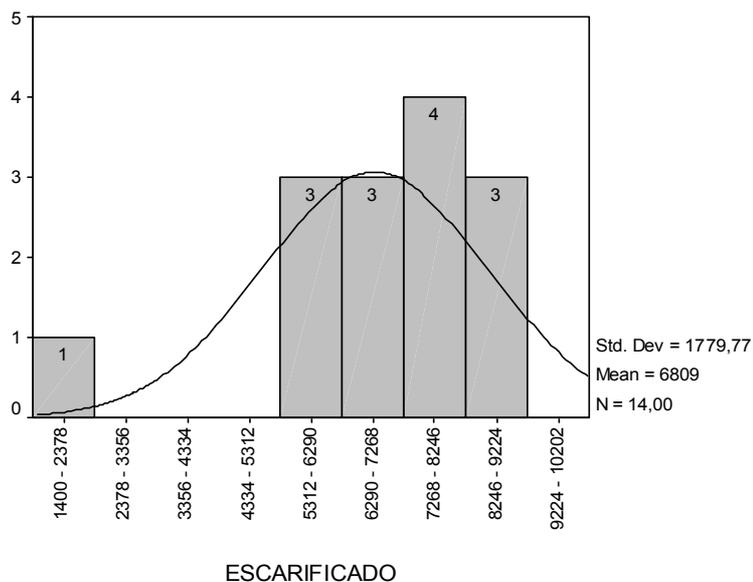
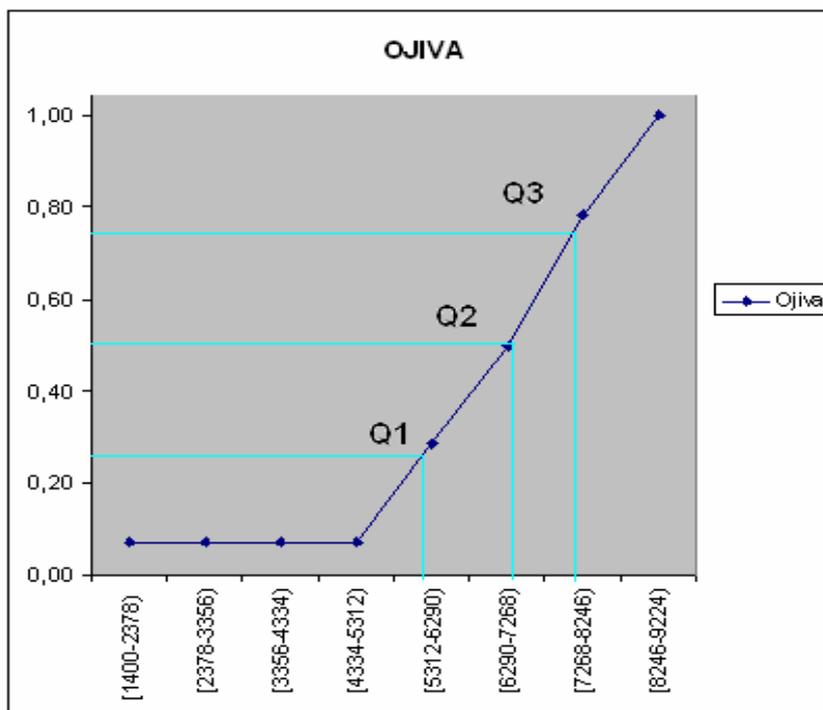


TABLA XXVII: TABLA DE FRECUENCIA DE ESCARIFICADO

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[1400-2378)	1889	1	0,0714286	0,0714286
[2378-3356)	2867	0	0,0000000	0,0714286
[3356-4334)	3845	0	0,0000000	0,0714286
[4334-5312)	4823	0	0,0000000	0,0714286
[5312-6290)	5801	3	0,2142857	0,2857143
[6290-7268)	6779	3	0,2142857	0,5000000
[7268-8246)	7757	4	0,2857143	0,7857143
[8246-9224)	8735	3	0,2142857	1,0000000

GRAFICO XXXI: OJIVA DE ESCARIFICADO



El primer cuartil es 6228 has siendo el 25% del área ablandado, el segundo cuartil es 7305,53 has y el tercer cuartil es 7552,8has.

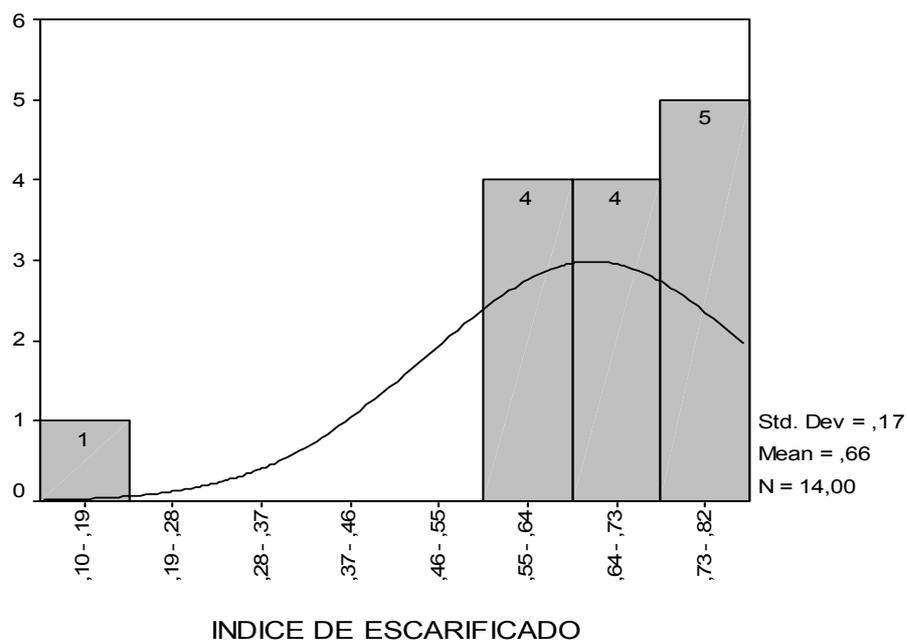
3.2.11 INDICADOR DE ESCLARIFICADO

**TABLA XXVIII: ESTADÍSTICA DESCRIPTIVA: VARIABLE
ESCARIFICADO MEDIANTE UN INDICADOR DE VECES**

N de Casos	14
Mínimo(1997)	0,13
Máximo(2002)	0,79
Rango	0,66
Mediana	0,71
Moda	0,72
Media	0,660714286
Nivel de confianza(95,0%)	0,096997763
Error estándar	0,044898717
Desviación estándar	0,167995617
Varianza	0,028222527
sesgo	-2,688396674

Fuente: Ingenio Valdez

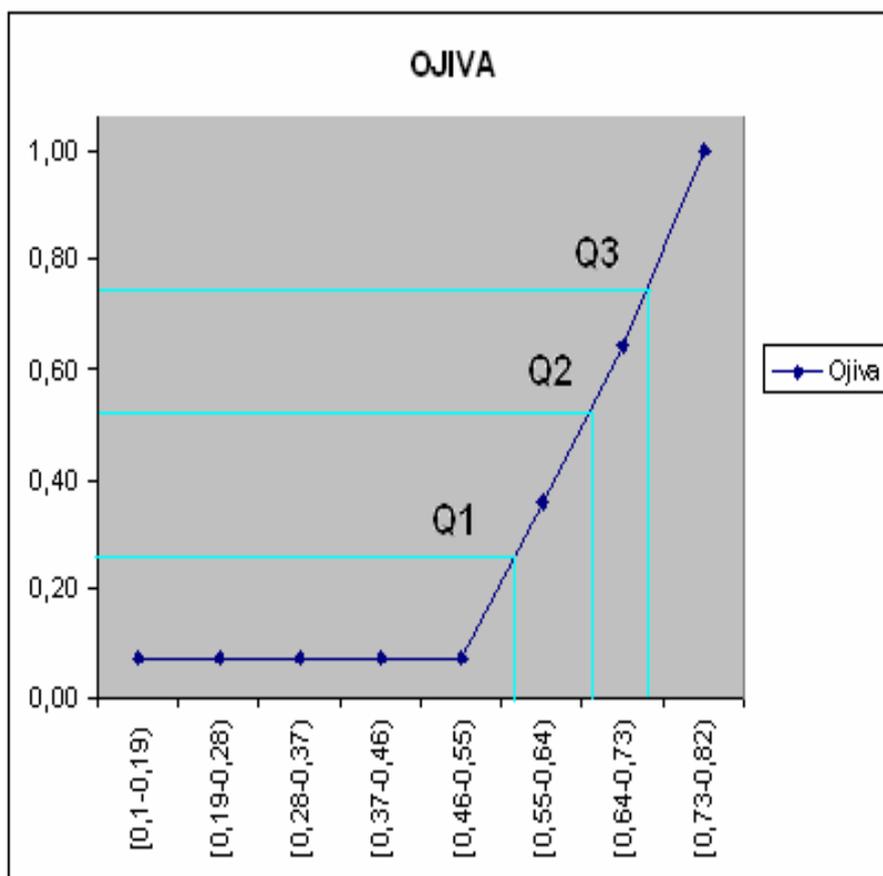
**GRAFICO XXXII: HISTOGRAMA DE FRECUENCIAS DE
ESCARIFICADO MEDIANTE UN INDICADOR DE VECES**



**TABLA XXIX: TABLA DE FRECUENCIA DE ESCARIFICADO
MEDIANTE UN INDICADOR DE VECES**

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[0,1-0,19)	0,145	1	0,0714286	0,0714286
[0,19-0,28)	0,235	0	0,0000000	0,0714286
[0,28-0,37)	0,325	0	0,0000000	0,0714286
[0,37-0,46)	0,415	0	0,0000000	0,0714286
[0,46-0,55)	0,505	0	0,0000000	0,0714286
[0,55-0,64)	0,595	4	0,2857143	0,3571429
[0,64-0,73)	0,685	4	0,2857143	0,6428571
[0,73-0,82)	0,775	5	0,3571429	1,0000000

GRAFICO XXXIII: OJIVA DE ESCARIFICADO MEDIANTE UN INDICADOR DE VECES



La frecuencia acumulada de la variable rendimiento muestra que el 25% de los datos Q1 es menor o igual que 0.62 veces, el 50% Q2 es menor o igual que 0.71 aproximadamente 1 veces y el 75% Q3 es menor igual que 0.75 .

3.2.12 RIEGO

TABLA XXX: ESTADÍSTICA DESCRIPTIVA: VARIABLE RIEGO

N de Casos	14
Mínimo(1997)	621,89
Máximo(1996)	31369
Rango	30747,11
Mediana	25822,3
Media	24433,83857
Nivel de confianza(95,0%)	4304,127033
Error estándar	1992,311752
Desviación estándar	7454,547985
Varianza	55570285,65
sesgo	-2,764868714

Fuente: Ingenio Valdez

GRAFICO XXXIV: HISTOGRAMA DE FRECUENCIA DE RIEGO

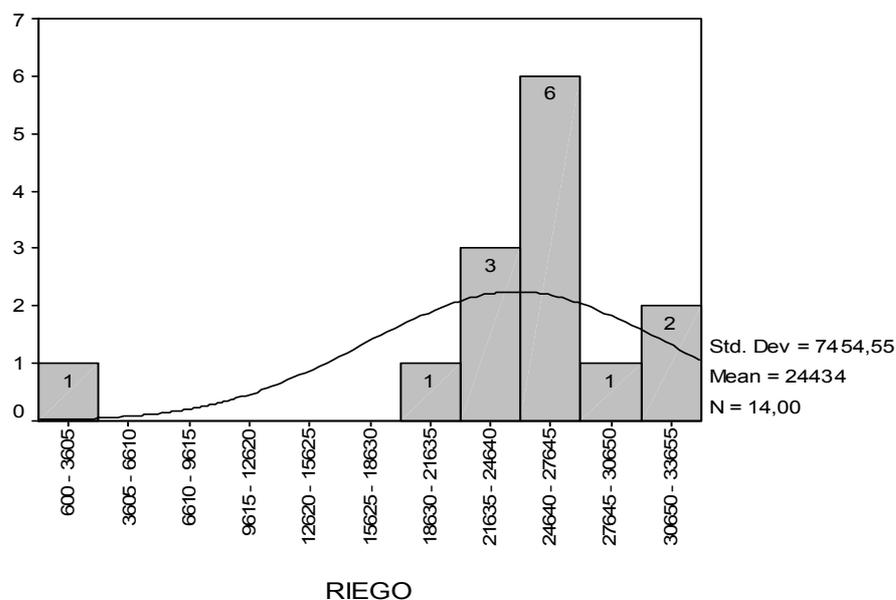
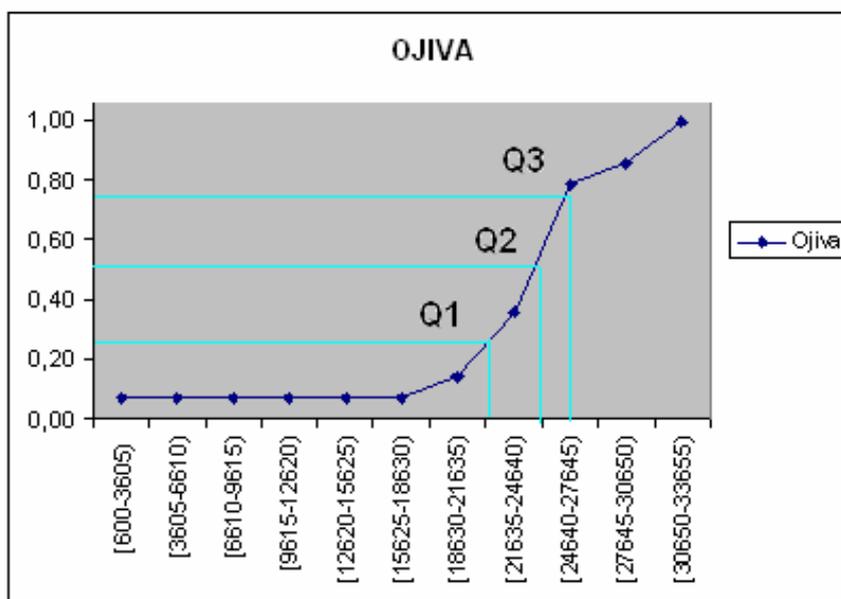


TABLA XXXI: TABLA DE FRECUENCIAS DE RIEGO

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[600-3605)	2102,5	1	0,0714286	0,0714286
[3605-6610)	5107,5	0	0,0000000	0,0714286
[6610-9615)	8112,5	0	0,0000000	0,0714286
[9615-12620)	11117,5	0	0,0000000	0,0714286
[12620-15625)	14122,5	0	0,0000000	0,0714286
[15625-18630)	17127,5	0	0,0000000	0,0714286
[18630-21635)	20132,5	1	0,0714286	0,1428571
[21635-24640)	23137,5	3	0,2142857	0,3571429
[24640-27645)	26142,5	6	0,4285714	0,7857143
[27645-30650)	29147,5	1	0,0714286	0,8571429
[30650-33655)	32152,5	2	0,1428571	1,0000000

GRAFICO XXXV: OJIVA DE RIEGO



El primer cuartil es 23828,24 el segundo cuartil o mediana es 25822.3, y el tercer cuartil es de 26959.3. El rango semicuartilico es $(Q3-Q1)/2$, por lo cual en este caso es de 1565,53.

3.2.13 INDICADOR DE RIEGO

TABLA XXXII: ESTADÍSTICA DESCRIPTIVA: VARIABLE RIEGO

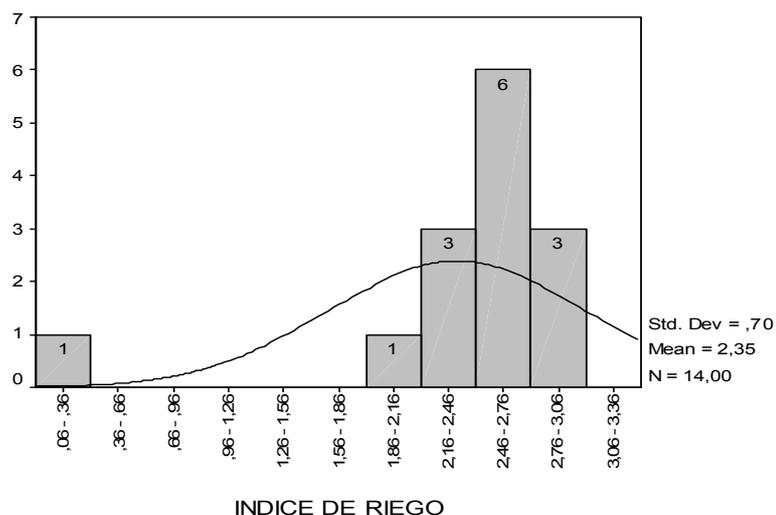
MEDIANTE UN INDICADOR DE VECES

N de Casos	14
Mínimo(1997)	0,06
Máximo(2002)	2,93
Rango	2,87
Mediana	2,49
Moda	2,5
Media	2,352857143
Nivel de confianza(95,0%)	0,403910387
Error estándar	0,186963676
Desviación estándar	0,699554018
Varianza	0,489375824
sesgo	-3,038052373

Fuente: Ingenio Valdez

GRAFICO XXXVI: HISTOGRAMA DE FRECUENCIA RIEGO

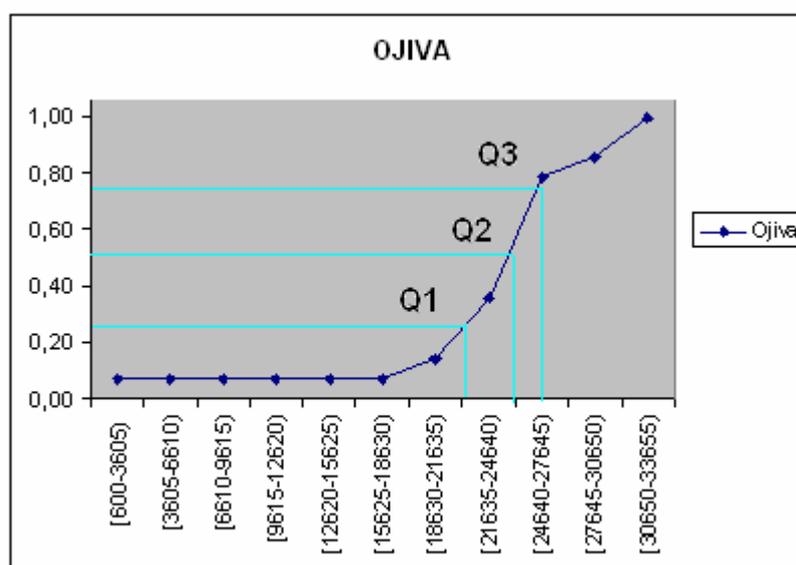
MEDIANTE UN INDICADOR DE VECES



**TABLA XXXIII: TABLA DE FRECUENCIAS RIEGO MEDIANTE
UN INDICADOR DE VECES**

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[0,057-0,357)	0,207	1	0,0714286	0,0714286
[0,357-0,657)	0,507	0	0,0000000	0,0714286
[0,657-0,957)	0,807	0	0,0000000	0,0714286
[0,957-1,257)	1,107	0	0,0000000	0,0714286
[1,257-1,557)	1,407	0	0,0000000	0,0714286
[1,557-1,857)	1,707	0	0,0000000	0,0714286
[1,857-2,157)	2,007	1	0,0714286	0,1428571
[2,157-2,457)	2,307	3	0,2142857	0,3571429
[2,457-2,757)	2,607	6	0,4285714	0,7857143
[2,757-3,057)	2,907	3	0,2142857	1,0000000

**GRAFICO XXXVII: OJIVA DE RIEGO MEDIANTE UN INDICADOR
DE VECES**



La frecuencia acumulada de la variable rendimiento muestra que el 25% de los datos Q1 es menor o igual que 2.4 mas de 2 veces, el 50% Q2 es menor o igual que 2.49 y el 75% Q3 es menor igual que 2.57 aproximadamente 3 veces.

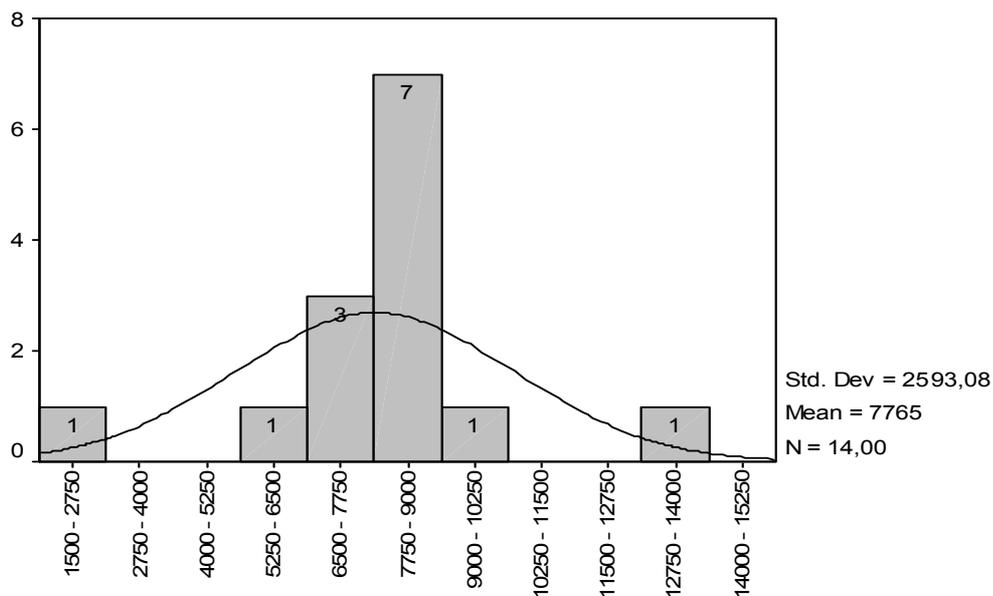
3.2.14 SACADA DE PAJA CAUCA

**TABLA XXXIV: ESTADÍSTICA DESCRIPTIVA: VARIABLE
SACADA DE PAJA CAUCA**

N de Casos	14
Mínimo(1997)	1613
Máximo(1996)	13769
Rango	12156
Mediana	7979,58
Media	7839,156429
Nivel de confianza(95,0%)	1464,461078
Error estándar	677,8756745
Desviación estándar	2536,378525
Varianza	6433216,022
sesgo	-0,177263162

Fuente: Ingenio Valdez

**GRAFICO XXXVIII: HISTOGRAMA DE FRECUENCIAS SACADA
DE PAJA CAUCA**

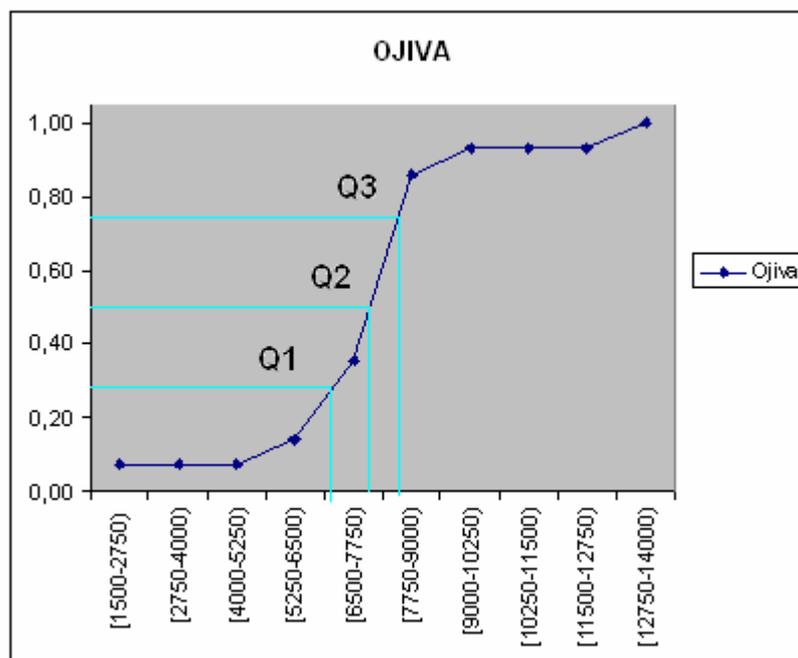


SACADA DE PAJA CAUCA

**TABLA XXXV: TABLA DE FRECUENCIAS SACADA DE PAJA
CAUCA**

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[1500-2750)	2125	1	0,0714286	0,0714286
[2750-4000)	3375	0	0,0000000	0,0714286
[4000-5250)	4625	0	0,0000000	0,0714286
[5250-6500)	5875	1	0,0714286	0,1428571
[6500-7750)	7125	3	0,2142857	0,3571429
[7750-9000)	8375	7	0,5000000	0,8571429
[9000-10250)	9625	1	0,0714286	0,9285714
[10250-11500)	10875	0	0,0000000	0,9285714
[11500-12750)	12125	0	0,0000000	0,9285714
[12750-14000)	13375	1	0,0714286	1,0000000

GRAFICO XXXIX: OJIVA DE SACADA DE PAJA CAUCA



La frecuencia acumulada de la variable roza de cantero muestra que el 25% de los datos Q1 es menor e igual 6686 hás donde Q2 es 7979.58 que es igual a su media y Q3 es 8187 siendo el valor que representa el 75% de la roza de cantero.

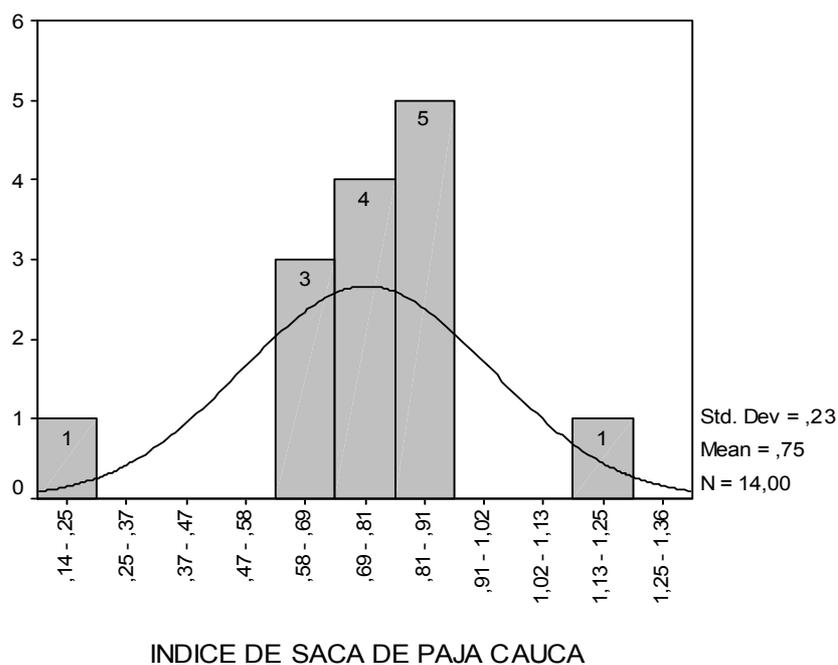
3.2.15 INDICADOR DE SACADA DE PAJA CAUCA

**TABLA XXXVI: ESTADÍSTICA DESCRIPTIVA: VARIABLE
SACADA DE PAJA CAUCA MEDIANTE UN INDICADOR DE
VECES**

N de Casos	14
Mínimo(1997)	0,15
Máximo(1996)	1,24
Rango	1,09
Mediana	0,775
Moda	0,81
Media	0,752857143
Nivel de confianza(95,0%)	0,132931747
Error estándar	0,061531986
Desviación estándar	0,230231608
Varianza	0,053006593
sesgo	-0,754838264

Fuente: Ingenio Valdez

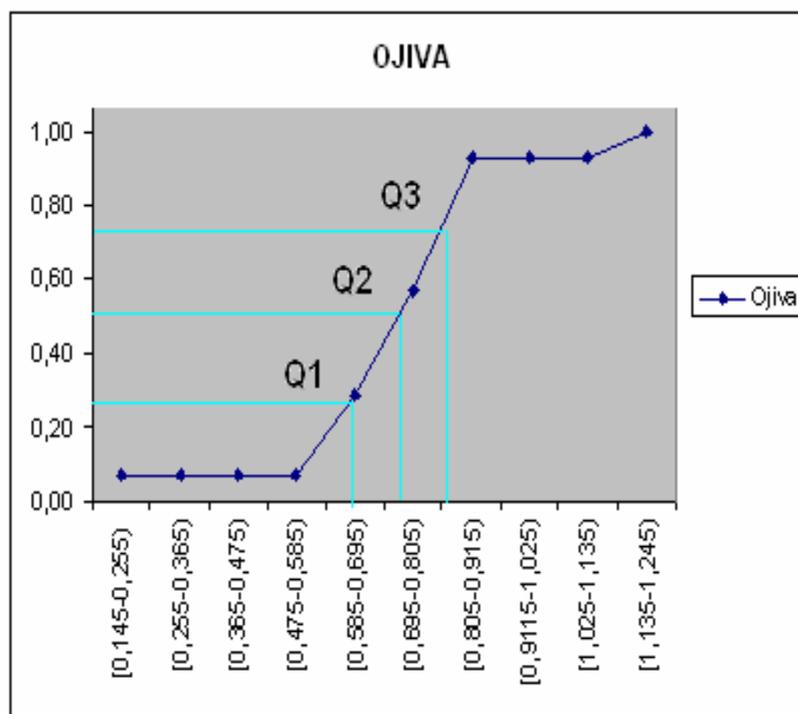
**GRAFICO XL: HISTOGRAMA DE FRECUENCIAS DE SACADA DE
PAJA CAUCA MEDIANTE UN INDICADOR DE VECES**



**TABLA XXXVII: TABLA DE FRECUENCIAS SACADA DE PAJA
CAUCA MEDIANTE UN INDICADOR DE VECES**

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[0,145-0,255)	0,2	1	0,0714286	0,0714286
[0,255-0,365)	0,31	0	0,0000000	0,0714286
[0,365-0,475)	0,42	0	0,0000000	0,0714286
[0,475-0,585)	0,53	0	0,0000000	0,0714286
[0,585-0,695)	0,64	3	0,2142857	0,2857143
[0,695-0,805)	0,75	4	0,2857143	0,5714286
[0,805-0,915)	0,86	5	0,3571429	0,9285714
[0,9115-1,025)	0,96825	0	0,0000000	0,9285714
[1,025-1,135)	1,08	0	0,0000000	0,9285714
[1,135-1,245)	1,19	1	0,0714286	1,0000000

**GRAFICO XLI: OJIVA SACADA DE PAJA CAUCA MEDIANTE UN
INDICADOR DE VECES**



La frecuencia acumulada de la variable rendimiento muestra que el 25% de los datos Q1 es menor o igual que 0,67, el 50% Q2 es menor o igual que 0,775 veces y el 75% Q3 es menor igual que 0,81 que es aproximadamente una vez.

3.2.16 APORQUE

**TABLA XXXVIII: ESTADÍSTICA DESCRIPTIVA: VARIABLE
APORQUE**

N de Casos	14
Mínimo(1999)	0
Máximo(2000)	5074
Rango	5074
Mediana	2691,5
Media	2538,345714
Nivel de confianza(95,0%)	900,1537905
Error estándar	416,666832
Desviación estándar	1559,02453
Varianza	2430557,485
sesgo	-0,215597715

Fuente: Ingenio Valdez

GRAFICO XLII: HISTOGRAMA DE FRECUENCIAS APORQUE

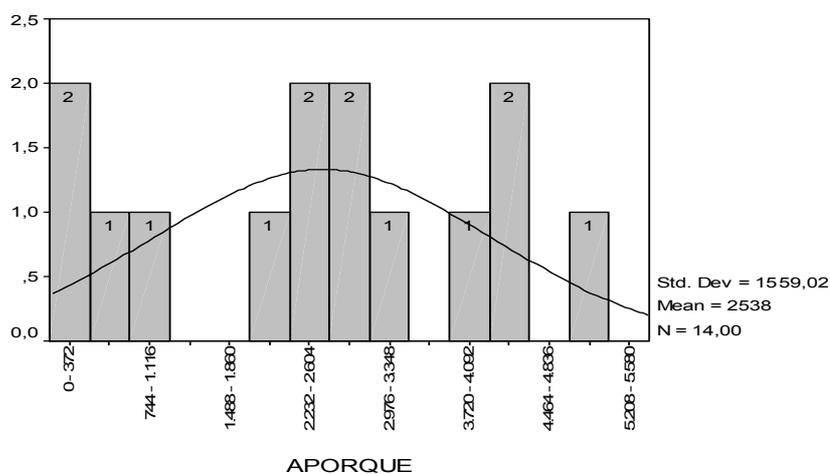
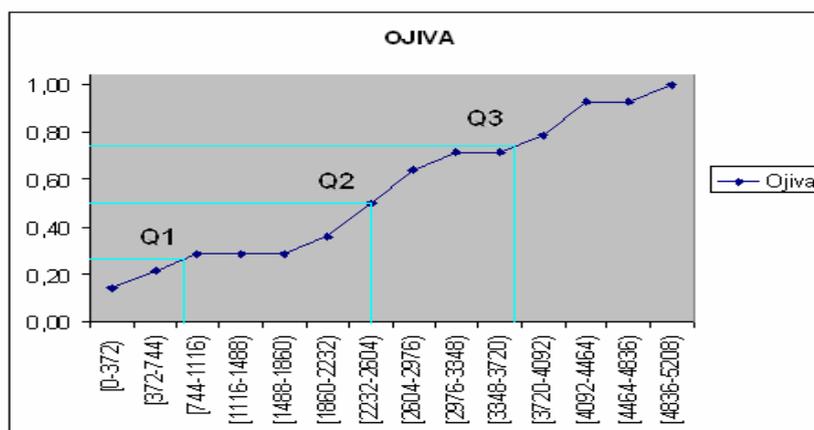


TABLA XXXIX : TABLA DE FRECUENCIAS DE APORQUE

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[0-372)	186	2	0,1428571	0,1428571
[372-744)	558	1	0,0714286	0,2142857
[744-1116)	930	1	0,0714286	0,2857143
[1116-1488)	1302	0	0,0000000	0,2857143
[1488-1860)	1674	0	0,0000000	0,2857143
[1860-2232)	2046	1	0,0714286	0,3571429
[2232-2604)	2418	2	0,1428571	0,5000000
[2604-2976)	2790	2	0,1428571	0,6428571
[2976-3348)	3162	1	0,0714286	0,7142857
[3348-3720)	3534	0	0,0000000	0,7142857
[3720-4092)	3906	1	0,0714286	0,7857143
[4092-4464)	4278	2	0,1428571	0,9285714
[4464-4836)	4650	0	0,0000000	0,9285714
[4836-5208)	5022	1	0,0714286	1,0000000

GRAFICO XLIII: OJIVA SACADA DE PAJA CAUCA MEDIANTE UN INDICADOR DE VECES



La frecuencia acumulada de la variable roza de cantero muestra que el 25% de los datos Q1 es menor e igual 944.1 has donde Q2 es 20584 que es igual a su media y Q3 es 2691.5 siendo el valor que representa el 75% de la roza de cantero cuyo valor es 3114.

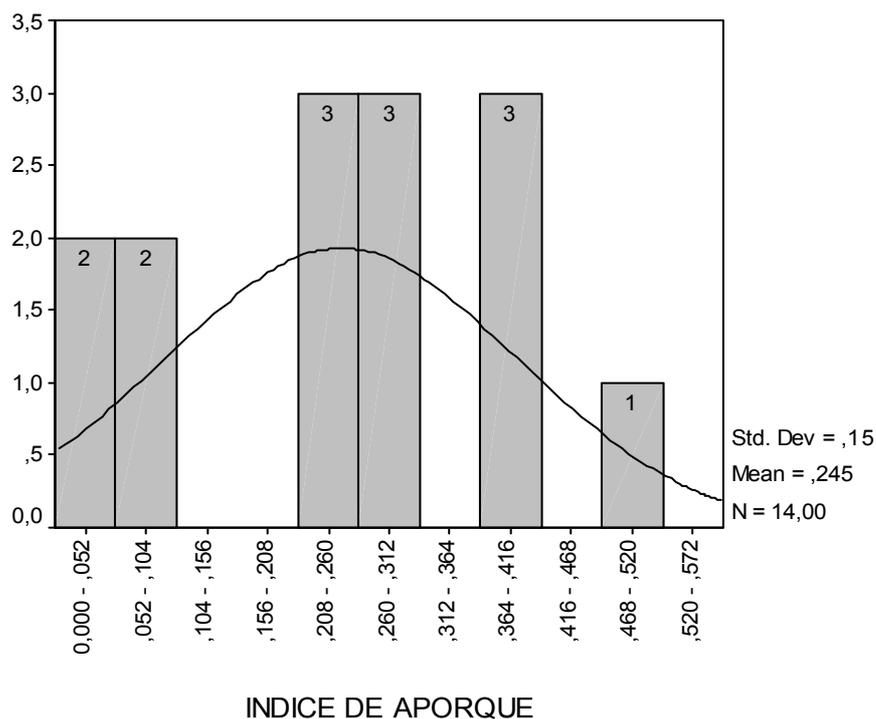
3.2.17 INDICADOR DE APORQUE

**TABLA XL: ESTADÍSTICA DESCRIPTIVA: VARIABLE
APORQUE MEDIANTE UN INDICADOR DE VECES**

N de Casos	14
Mínimo(1999)	0
Máximo(2001)	0,48
Rango	0,48
Mediana	0,26
Moda	0,24
Media	0,245
Nivel de confianza(95,0%)	0,086851339
Error estándar	0,040202099
Desviación estándar	0,150422482
Varianza	0,022626923
sesgo	-0,285320242

Fuente: Ingenio Valdez

**GRAFICO XLIV: HISTOGRAMA DE FRECUENCIAS APORQUE
MEDIANTE UN INDICADOR DE VECES**

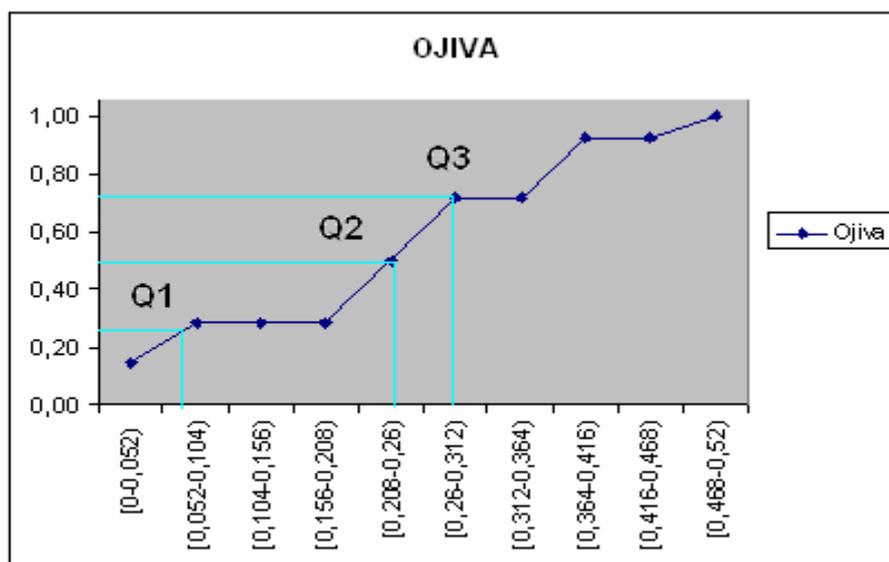


**TABLA XLI: TABLA DE FRECUENCIAS APORQUE MEDIANTE
UN INDICADOR DE VECES**

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[0-0,052)	0,026	2	0,1428	0,1428
[0,052-0,104)	0,078	2	0,1428	0,2857
[0,104-0,156)	0,13	0	0,0000	0,2857
[0,156-0,208)	0,182	0	0,0000	0,2857
[0,208-0,26)	0,234	3	0,2142	0,5000

[0,26-0,312)	0,286	3	0,2142	0,7142
[0,312-0,364)	0,338	0	0,0000	0,7142
[0,364-0,416)	0,39	3	0,2142	0,9285
[0,416-0,468)	0,442	0	0,0000	0,9285
[0,468-0,52)	0,494	1	0,0714	1,0000

GRAFICO XLV: OJIVA DE APORQUE MEDIANTE UN INDICADOR DE VECES



La frecuencia acumulada de la variable rendimiento muestra que el 25% de los datos Q1 es menor o igual que 0,09, el 50% Q2 es menor o igual que 0,26 y el 75% Q3 es menor igual que 0,3, es mínima la cantidad de tierra que se acumula alrededor de las plantas.

3.2.18 Fertilización

**TABLA XLII: ESTADÍSTICA DESCRIPTIVA: VARIABLE
FERTILIZACIÓN**

N de Casos	14
Mínimo(1997)	936
Máximo(1996)	11 150
Rango	10 214
Mediana	8 264,895
Media	8 146,766 429
Nivel de confianza(95,0%)	1 343,158 157
Error estándar	6 21,726 4871
Desviación estándar	2 326,287 503
Varianza	5 411 613,547
sesgo	-2,380 618 708

Fuente: Ingenio Valdez

**GRAFICO XLVI: HISTOGRAMA DE FRECUENCIAS DE
FERTILIZACIÓN**

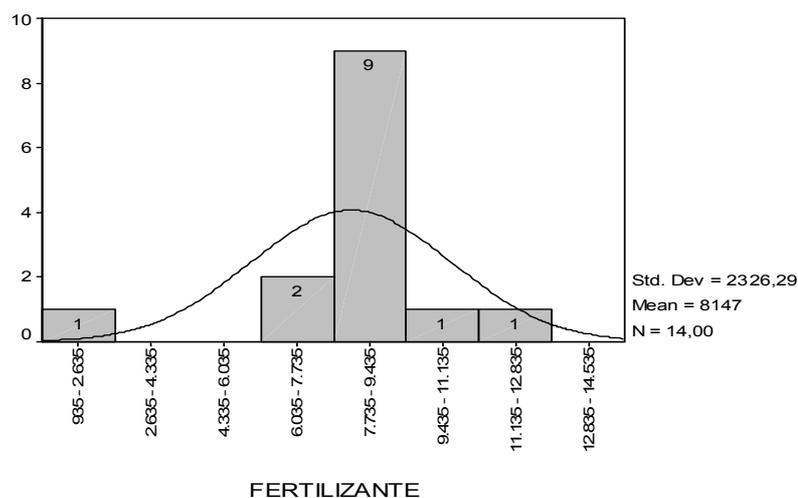
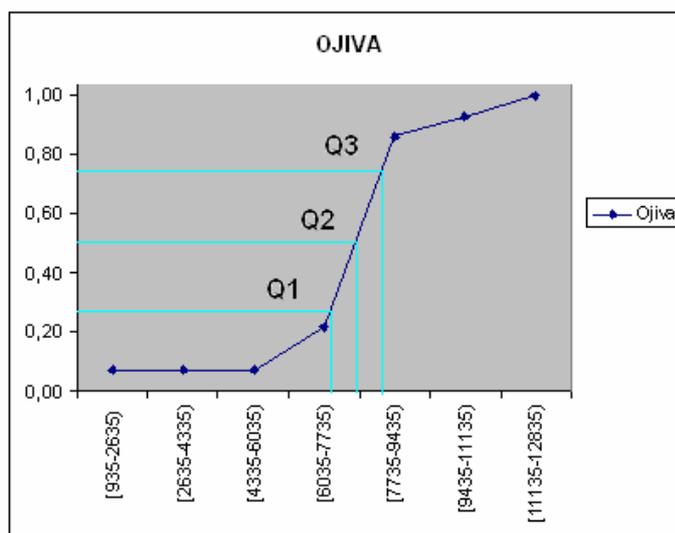


TABLA XLIII: TABLA DE FRECUENCIAS DE FERTILIZACIÓN

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[935-2635)	1785	1	0,0714286	0,0714286
[2635-4335)	3485	0	0,0000000	0,0714286
[4335-6035)	5185	0	0,0000000	0,0714286
[6035-7735)	6885	2	0,1428571	0,2142857
[7735-9435)	8585	9	0,6428571	0,8571429
[9435-11135)	10285	1	0,0714286	0,9285714
[11135-12835)	11985	1	0,0714286	1,0000000

GRAFICO XLVII: OJIVA DE FERTILIZACIÓN



La frecuencia acumulada de la variable roza de cantero muestra que el 25% de los datos Q1 es menor e igual 7793 hás donde Q2 es

8264.89 que es igual a su media y Q3 es 8926.8 siendo el valor que representa el 75% de la roza de cantero.

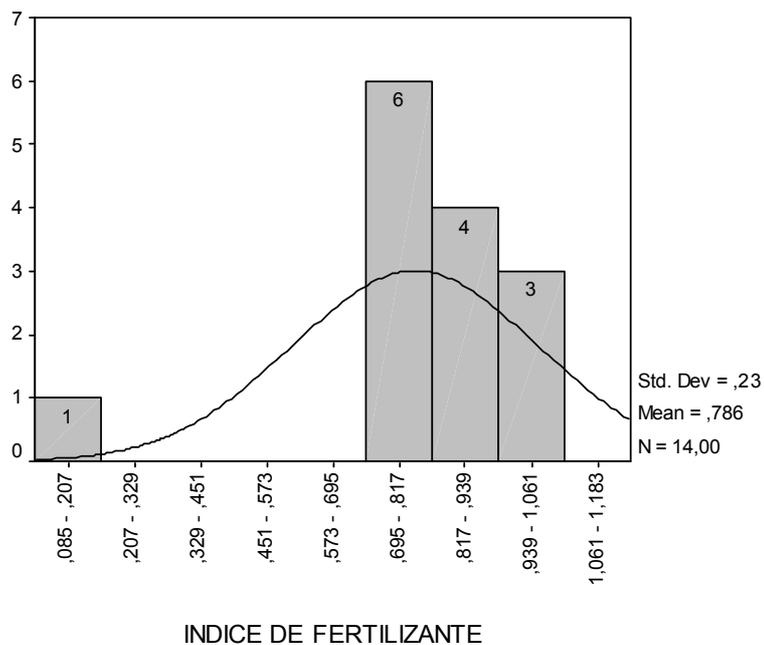
3.2.19 INDICADOR DE FERLIZACIÓN

**TABLA XLIV: ESTADÍSTICA DESCRIPTIVA: VARIABLE
FERTILIZACIÓN MEDIANTE UN INDICADOR DE VECES**

N de Casos	14
Mínimo(1997)	0,09
Máximo(1996)	1
Rango	0,91
Mediana	0,825
Moda	0,99
Media	0,785714286
Nivel de confianza(95,0%)	0,130523354
Error estándar	0,060417179
Desviación estándar	0,226060383
Varianza	0,051103297
sesgo	-2,393958846

Fuente: Ingenio Valdez

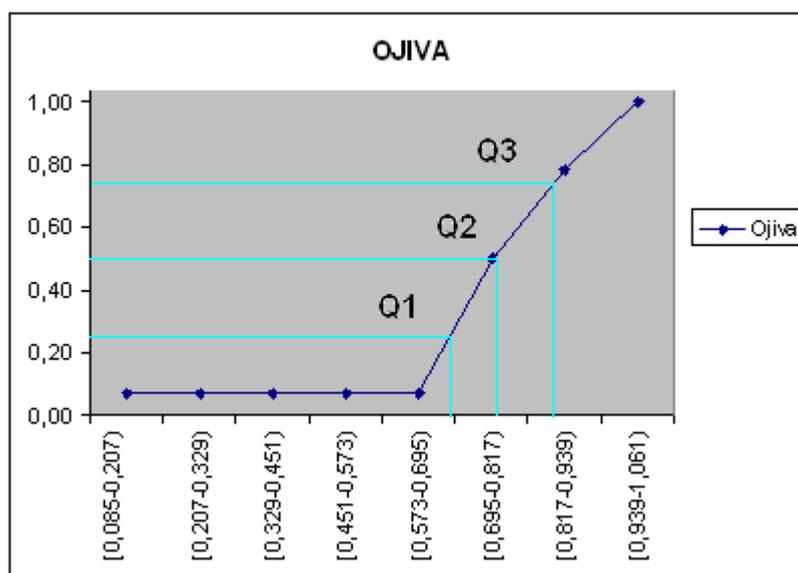
**GRAFICO XLVIII: HISTOGRAMA DE FRECUENCIAS
FERTILIZACIÓN MEDIANTE UN INDICADOR DE VECES**



**TABLA XLV: TABLA DE FRECUENCIAS FERTILIZACIÓN
MEDIANTE UN INDICADOR DE VECES**

Intervalo de clase	Marca de clase	Frecuencia Absoluta	Frecuencia relativa	Frecuencia acumulada relativa
[0,085-0,207)	0,146	1	0,0714	0,0714
[0,207-0,329)	0,268	0	0,0000	0,0714
[0,329-0,451)	0,39	0	0,0000	0,0714
[0,451-0,573)	0,512	0	0,0000	0,0714
[0,573-0,695)	0,634	0	0,0000	0,0714
[0,695-0,817)	0,756	6	0,4285	0,5000
[0,817-0,939)	0,878	4	0,2857	0,7857
[0,939-1,061)	1	3	0,2142	1,0000

GRAFICO XLIX: OJIVA DE FERTILIZACIÓN MEDIANTE UN INDICADOR DE VECES



La frecuencia acumulada de la variable rendimiento muestra que el 25% de los datos Q1 es menor o igual que 0.74, el 50% Q2 es menor o igual que 0.825 y el 75% Q3 es menor igual que 0.88.

3.3 ANALISIS MULTIVARIADO DE LAS VARIABLES DE INVESTIGACIÓN

3.3.1 Análisis de Componentes Principales

Con el objetivo de medir las variables que influyen en la producción de caña de azúcar tomando como referencia el año 1989 hasta el año 2002 uno de los métodos mas adecuados es el análisis estadístico multivariado, permitirá extraer la máxima información posible del conjunto de datos y expresarlo en un número reducido de combinaciones lineales, para las cuales la varianza no explicada fuera mínima y estudiar la relación lineal entre las variables que influyen en la producción de caña de azúcar.

Para así lograr obtener las variables de mayor importancia dentro del cultivo de la caña de azúcar y poder obtener mejores resultados del análisis realizado, el objetivo más importante de este análisis es reducir en lo máximo las variables y obtener las más representativas.

TABLA XLVI: NOMINA DE VARIABLES DE INTERÉS

VARIABLES	REPRESENTACIÓN	UNIDADES
Superficie sembrada	X_1	Has
Superficie cosechada	X_2	Has
Producción	X_3	Tons
Roza de canteros	X_4	Has
Herbicidas	X_5	Has
Rastrillo	X_6	Has
Escarificado	X_7	Has
Riego	X_8	Has
Sacada de paja cauca	X_9	Has
Aporque	X_{10}	Has
Fertilizante	X_{11}	Has

La matriz de varianzas y covarianzas S muestra que los valores de las variables Producción, Riego, Roza son mayores que las otras variables. Las covarianzas asumen valores del mismo orden de magnitud que las varianzas, lo que indica una dependencia entre las variables.

La variable producción está dada en Toneladas (tons). Lo contrario de la variables Riego y Roza están dadas en Hectáreas (Has), son las que tienen una mayor varianza y covarianza esto se debe a que los valores de sus magnitudes son mayores,

A continuación se presentan en la tabla de los valores propios y la proporción de la variación total explicada por cada componente al utilizar la matriz S.

TABLA XLVII: VALORES PROPIOS Y PROPORCIÓN DE LA VARIANZA TOTAL EXPLICADA DE LOS DATOS

Componente	Valor Propio	Porcentaje de explicación	Porcentaje acumulado
1	13077025762,307	99,903	99,903
2	8173631,103	0,062	99,965
3	2288588,949	0,017	99,983
4	996594,515	0,008	99,990
5	852757,420	0,007	99,997
6	259397,221	0,002	99,999
7	89778,147	0,001	100,000
8	36487,022	0,000	100,000
9	17413,405	0,000	100,000
10	2076,127	0,000	100,000
11	566,226	0,000	100,000

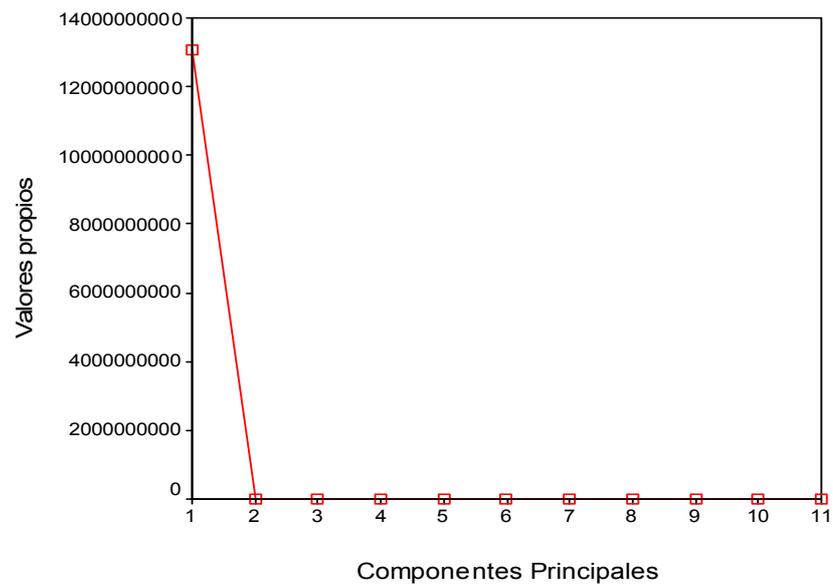
Cada valor propio tiene una componente asociada en total tenemos 11 componente principales, como podemos observar la primera componente explica un 99.90% de la variabilidad total y la segunda resume el 0.06% es decir que la combinación lineal de las variables originales con los dos primeros componentes sintetizan el 99.96% de la variación total del conjunto de datos.

Pero para poder reemplazar las 11 variables por los dos primeros componentes hay que examinar también los valores propios y los coeficientes de correlación de las nuevas variables con las originales

TABLA XLVIII: VECTORES PROPIOS

VECTORES PROPIOS		
Variable	1*	2*
superficie sembrada	0,0011	-0,0497
superficie cosechada	0,0147	0,1268
producción	0,9954	-0,0742
roza de cantero	0,0496	0,5032
herbicida	0,0330	-0,1266
rastrillo	0,0135	0,0538
escarificado	0,0149	0,1365
riego	0,0620	0,7825
sacada de paja cauca	0,0213	-0,0866
aporque	0,0122	-0,1631
fertilización	0,0194	0,1969

Se ha escogido 2 componentes, la tabla anterior muestra que el primer componente muestra un coeficiente de 0.995 para la producción y 0.062 para la variable riego, es decir que la primera combinación lineal refleja la variación en producción y la segunda combinación lineal refleja el riego con un coeficiente de 0.78 y la roza con un coeficiente de 0.50.

GRAFICO L: COMPONENTES PRINCIPALES

Los componentes principales quedan expresados de la siguiente manera, los cuales son la combinación lineal de los vectores propios de la matriz con las variables.

$$Y_1 = 0.0011X_1 + 0.0147X_2 + 0.9954X_3 + 0.0496X_4 + 0.0330X_5 + 0.0135X_6 + 0.0149X_7 + 0.0620X_8 + 0.0213X_9 + 0.0122X_{10} + 0.0194X_{11}$$

$$Y_2 = -0.0497X_1 + 0.1268X_2 + -0.0742X_3 + 0.5032X_4 + -0.1266X_5 + 0.0538X_6 + 0.1365X_7 + 0.7825X_8 + 0. -0.0866X_9 + -0.11631X_{10} + 0.1969X_{11}$$

Debido a que las varianzas son diferentes entre sí ya que las variables se encuentran en distintas unidades es recomendable estandarizar los datos originales a fin de homogeneizar las magnitudes para un mejor estudio de los componentes principales.

A partir de la matriz de varianzas y covarianzas S se obtiene la matriz de correlación R.

TABLA XLIX: MATRIZ DE CORRELACIÓN DE LOS DATOS

MATRIZ DE CORRELACIÓN											
	APORQUE	COSECHA	ESCARIF	FERTILIZACIÓN	HERBICIDA	PRODUCCIÓN	RASTRILLO	RIEGO	ROZA	SACADA PAJA CAUCA	SIEMBRA
Aporque	1,0000	0,8376	0,8228	0,7774	0,9630	0,8988	0,8833	0,7597	0,8108	0,8817	0,1955
Cosecha	0,8376	1,0000	0,9952	0,9795	0,9284	0,9673	0,9773	0,9819	0,9776	0,8987	0,1962
Escarificado	0,8228	0,9952	1,0000	0,9813	0,9166	0,9618	0,9788	0,9840	0,9652	0,8922	0,2233
Fertilización	0,7774	0,9795	0,9813	1,0000	0,8766	0,9517	0,9346	0,9872	0,9430	0,9152	0,1379
Herbicida	0,9630	0,9284	0,9166	0,8766	1,0000	0,9622	0,9405	0,8795	0,9271	0,9078	0,1564
Producción	0,8988	0,9673	0,9618	0,9517	0,9622	1,0000	0,9503	0,9508	0,9535	0,9423	0,1537
Rastrillo	0,8833	0,9773	0,9788	0,9346	0,9405	0,9503	1,0000	0,9303	0,9366	0,8752	0,3406
Riego	0,7597	0,9819	0,9840	0,9872	0,8795	0,9508	0,9303	1,0000	0,9686	0,8786	0,1025
Roza	0,8108	0,9776	0,9652	0,9430	0,9271	0,9535	0,9366	0,9686	1,0000	0,8538	0,0695
Sacada de paja ca	0,8817	0,8987	0,8922	0,9152	0,9078	0,9423	0,8752	0,8786	0,8538	1,0000	0,1089
Siembra	0,1955	0,1962	0,2233	0,1379	0,1564	0,1537	0,3406	0,1025	0,0695	0,1089	1,0000

Observando y analizando la matriz de correlación se observa la relación lineal entre las variables, la mayor correlación positiva la tiene escarificado y cosecha con un coeficiente de 0.99, esto se debe a su poca diferencia de magnitudes entre sus unidades, fertilización y riego también tienen una fuerte relación lineal positiva con un coeficiente de 0.9872, es decir que tanto el riego como la

fertilización son importante para el desarrollo de caña de azúcar, cosecha y la producción también tienen una fuerte correlación con un coeficiente de 0.967, es decir que al aumentar la producción obviamente aumenta la cosecha y viceversa.

Al observar la Matriz de correlación podemos darnos cuenta que son pocos los datos que tienen una débil correlación entre ellos, la mayoría están altamente correlacionados entre sí.

Siembra y cosecha tiene una correlación débilmente positiva con un coeficiente de 0.19 es decir que no necesariamente si aumenta la siembra aumenta la cosecha esto es debido a que las hectáreas sembradas pueden ser cosechadas más de una vez ya que la caña de azúcar se le puede realizar hasta tres cosechas al mismo cultivo, Siembra y fertilización tiene una correlación débilmente positiva con un coeficiente de 0.13 esto se debe a que las hectáreas sembradas no siempre van a ser fertilizadas por que puede ocurrir que alguna no necesite ser fertilizada o porque esta ya fue fertilizada antes en su primera cosecha o también por que está a punto de ser cosechada esto suele suceder por que no siempre las hectáreas sembradas se encuentran en las mismas condiciones pueden ser unas cosechadas antes y otras después .

Se observa también una débil correlación entre roza y siembra con un coeficiente de 0.06 esto se debe a que las hectáreas sembradas

no van hacer rozadas si no es necesario y además la roza es aplicada cuando la mata de caña está de una altura considerable se puede decir durante su crecimiento.

Sacada de paja cauca y siembra tienen una correlación débilmente positiva con un coeficiente de 0.10 debido a que el sacado de paja cauca es una labor de controlar el crecimiento de la maleza por lo tanto esto no se va a dar durante la siembra sino mas bien mucho tiempo después durante su crecimiento.

Siembra y producción tienen una correlación débil mente positiva con un coeficiente de 0.1564 esto se debe a que no necesariamente si se siembra unas hectáreas estas van a producir una cantidad proporcional a las sembradas por que puede darse el caso que la producción sea alta como puede darse el caso que la producción sea baja ya que no depende de la cantidad sembrada sino mas bien de otras variables, estas son si las hectáreas fueron fertilizadas adecuadamente o regada u otras, o también si estas hectáreas se cosechan por primera vez o son cosechadas por segunda vez.

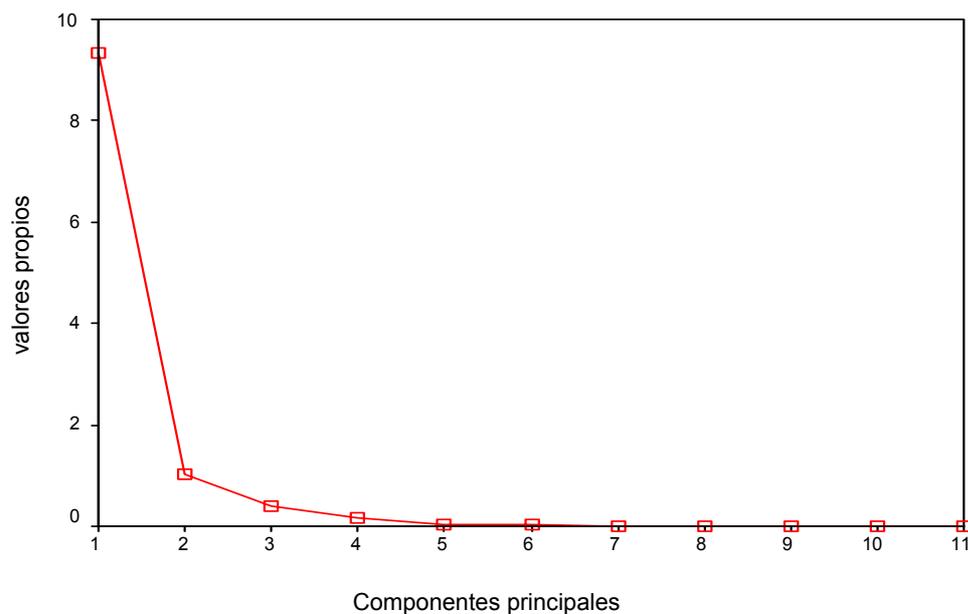
Siembra y herbicida están débilmente correlacionadas con un coeficiente de 0.15 debido a que cuando se realiza la siembra esta no es fertilizada sino más bien durante su crecimiento.

A partir de la matriz de correlación R se generarán los valores propios correspondientes y su proporción de varianza total explicada.

**TABLA L: PROPORCIÓN DE VARIANZA EXPLICADA (DATOS
STANDARIZADOS)**

Datos estandarizados: proporción de la varianza explicada			
componente	valor propio	porcentaje de explicación	porcentaje acumulado
1	9,3388	84,899	84,899
2	1,0232	9,302	94,201
3	0,4104	3,731	97,932
4	0,1553	1,412	99,344
5	0,0349	0,318	99,662
6	0,0184	0,167	99,829
7	0,0085	0,077	99,906
8	0,0057	0,052	99,958
9	0,0039	0,036	99,994
10	0,0005	0,005	99,999
11	0,0001	0,001	100

Las componentes se seleccionan dependiendo de tres criterios; que el valor característico sea mayor que la unidad, que el porcentaje de explicación acumulado determine que existe suficiente información para determinar la componente y el criterio del Codo el cual en la grafica de los valores característicos presenta un punto de inflexión que demuestra que alrededor de este existe un porcentaje de explicación suficiente para la componente.

GRAFICO LI: VALORES PROPIOS DATOS ESTANDARIZADOS

Los 2 primeros valores característicos son mayores que la unidad, poseen mayor explicación, el resto brinda una mínima contribución para la determinación de los componentes principales, ya que la suma de los 2 componentes alcanza el 94.20% de la variabilidad total explicada, y como se observa en la figura a partir del segundo hay un quebramiento; bajo estos criterios de las 11 componentes principales se deben escoger los 2 primeros lo cual significa que de una matriz de dimensión R^{11} se ha simplificado a una de R^2 .

A continuación presentamos los vectores propios u ortogonales asociados a los 2 componentes, los cuales son la combinación lineal entre las 11 variables de entrada y el vector ortogonal asociado a la matriz R. Además, los coeficientes de cada variable indican el grado de influencia de la variable en la componente principal.

**TABLA LI: VECTORES PROPIOS DE LOS DATOS
STANDARIZADOS**

VECTORES PROPIOS		
Variable	1*	2*
superficie sembrada	0,0638	0,9676
superficie cosechada	0,3242	-0,0006
producción	0,3235	-0,0408
roza de cantero	0,3163	-0,1239
herbicida	0,3152	-0,0202
rastrillo	0,3204	0,1557
escarificado	0,3228	0,0268
riego	0,3161	-0,0979
sacada de paja cauca	0,3062	-0,0768
aporque	0,2926	0,0441
fertilización	0,3172	-0,0623

Los componentes principales quedan expresados de la siguiente

manera

$$\begin{aligned}
 Y_1 = & 0.0638X_1 + 0.3242X_2 + 0.3235X_3 + 0.3163X_4 \\
 & + 0.3152X_5 + 0.3204X_6 + 0.3228X_7 + 0.3161X_8 \\
 & + 0.3062X_9 + 0.2926X_{10} + 0.3172X_{11}
 \end{aligned}$$

$$\begin{aligned}
 Y_1 = & 0.9676X_1 - 0.0006X_2 - 0.0408X_3 - 0.1239X_4 \\
 & - 0.0202X_5 + 0.1557X_6 + 0.0268X_7 - 0.0979X_8 \\
 & - 0.0768X_9 + 0.0441X_{10} - 0.0623X_{11}
 \end{aligned}$$

Interpretación de las componentes principales

Primer Componente

El primer componente es una combinación lineal de las variables con ponderación relativa mayor para las variables superficie cosechada, producción, escarificado, rastrillo, fertilización, roza de cantero, riego, es decir que la producción de caña de azúcar con valores elevados del primer componente estará asociada a valores elevados de dichas variables. Este componente se lo puede llamar **“Eje de producción cosechada”**.

Segundo Componente

El segundo componente tiene un valor alto con coeficiente positivo 0.96 para superficie sembrada, y tiene coeficiente negativo -0.12 para roza de cantero, -0.09 riego, -0.07 sacada de paja cauca, -0.06 fertilización, y con una contribución mínima de -0.0006 de superficie cosechada. Esto contribuye un contraste entre las variables de tal manera que los valores mas elevados del segundo componente principal serán aquellos que tengan una mayor cantidad de

superficie sembrada y menor roza de cantero, menor herbicida, menor fertilización, tendremos una menor superficie cosechada y menor producción.

En el cultivo de caña de azúcar nos encontramos con la situación que a menores insumos como es fertilización y otros tendremos una menor producción de caña, porque al aplicarle una cantidad no adecuada de insumos a la mata de caña este se debilitara y no servirá para una segunda o hasta una tercera cosecha a la misma mata. Se lo puede interpretar como **“Eje de labores de insumos”**.

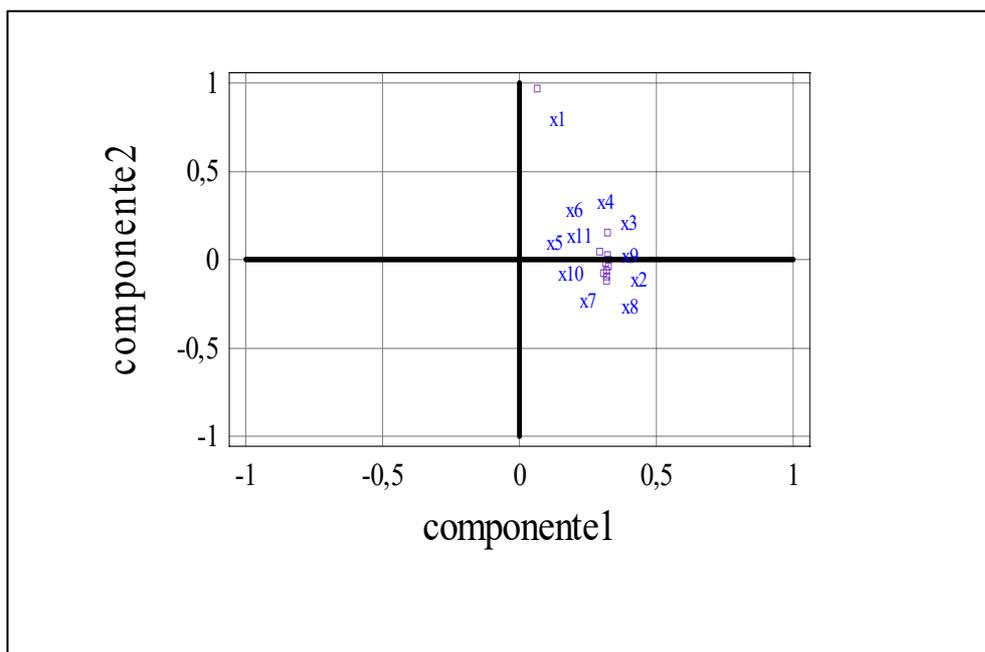
En la siguiente tabla se encuentra la correlación de las variables originales con cada una de las componentes, mientras mayor sea la correlación mayor carga tendrán en los componentes.

**TABLA LII: MATRIZ DE CARGA DE LOS DATOS
STANDARIZADOS**

MATRIZ DE CARGA		
Variable	F1	F2
superficie sembrada	0,1951	0,9788
superficie cosechada	0,9906	-0,0006
producción	0,9885	-0,0413
roza de cantero	0,9666	-0,1253
herbicida	0,9633	-0,0204
rastrillo	0,9791	0,1575
escarificado	0,9866	0,0272
riego	0,9661	-0,0992
sacada de paja cauca	0,9359	-0,0777
aporque	0,8942	0,0446
fertilización	0,9692	-0,0630

Podemos graficar la correlación $r(jk)$ de cada variable original con los dos primeros componentes. Las variables que más se acerquen al círculo de radio unitario que se encuentra en la escala -1 y 1 más influencia tendrán en los factores, en tanto que aquellas que se ubican cerca del origen de coordenadas son las que menos intervienen en los mismos.

GRAFICO LII: COMPONENTES EN EL PLANO



Como se puede observar en el Gráfico LII las variables se encuentran fuertemente correlacionadas con el primer y segundo componente

Existen variables que tienen alta correlación simultáneamente en algunos componentes. Para facilitar la interpretación se realiza lo que se denomina rotaciones de factores, que consiste en redistribuir la varianza explicada por cada factor. Para obtener este propósito se utiliza el método de **VARIMAX**, que consiste en minimizar el número de variables que tienen alta carga en un factor con lo cual simplificaríamos la información.

Se puede mostrar en la siguiente tabla los nuevos valores propios con su respectivo porcentaje de variación explicada

**TABLA LIII: DATOS ESTANDARIZADOS: PROPORCIÓN DE LA
VARIANZA TOTAL EXPLICADA DE LOS FACTORES
ROTADOS**

componente	valor propio	porcentaje de explicación	porcentaje acumulado
1	9,3388	84,899	84,899
2	1,0232	9,302	94,201

Mostraremos a continuación la matriz de carga.

**TABLA LIV: MATRIZ DE CARGA DE LOS FACTORES
ROTADOS**

MATRIZ DE CARGA		
Variable	F1	F2
superficie sembrada	0,0782	0,9950
superficie cosechada	0,9837	0,1164
producción	0,9864	0,0757
roza de cantero	0,9747	-0,0103
herbicida	0,9589	0,0934
rastrillo	0,9536	0,2720
escarificado	0,9765	0,1435
riego	0,9710	0,0156
sacada de paja cauca	0,9385	0,0333
aporque	0,8827	0,1499
fertilización	0,9699	0,0518

Primer Factor

El primer factor tiene mayor correlación con las variables producción, superficie cosechada, escarificado, roza de cantero.

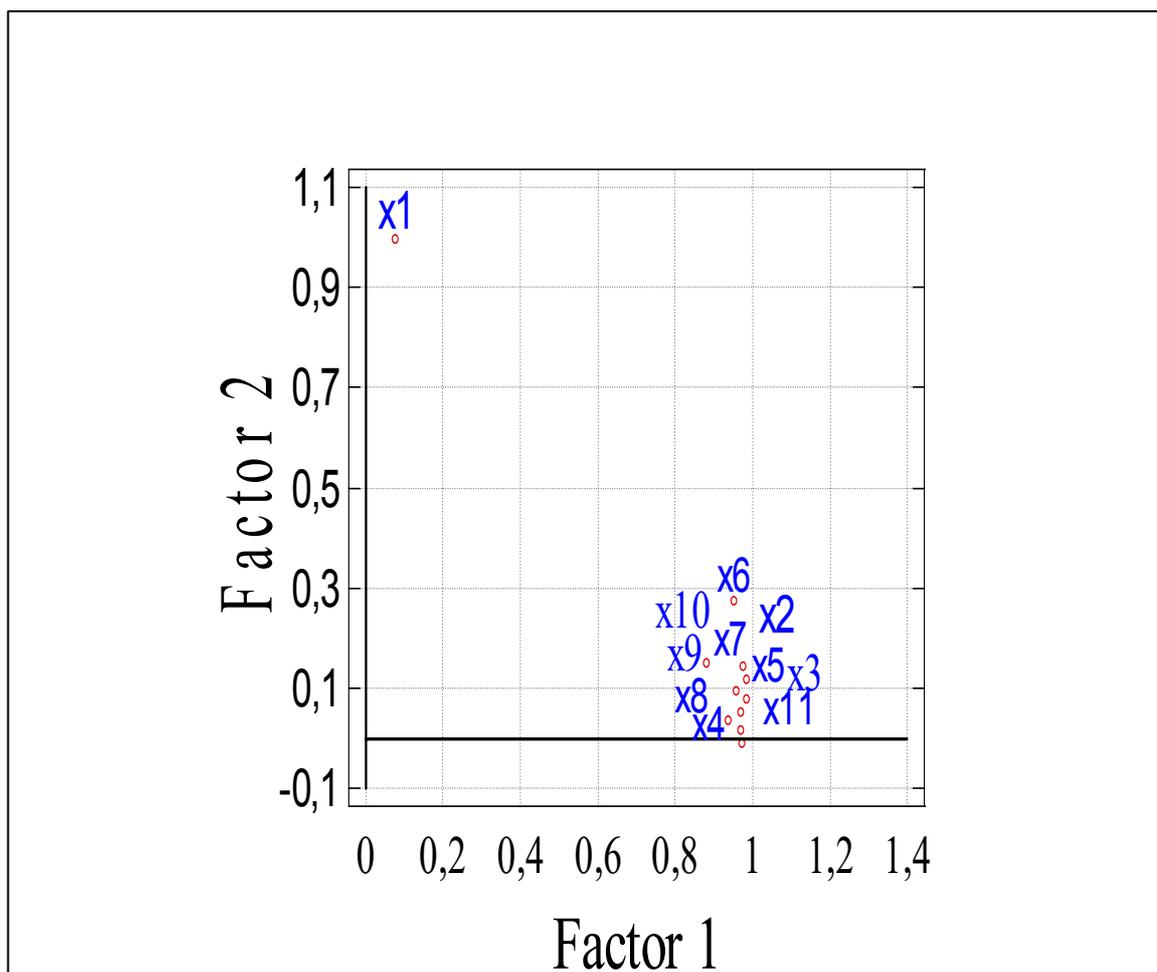
Podemos llamar a este factor “**Superficie agrícola de cosecha**”.

Segundo Factor

Este factor tiene mayor correlación con la variable superficie sembrada con un coeficiente de 0.99 y con un contraste con roza de

cantero con un coeficiente de -0.010 por lo cual a este factor lo llamaremos “**Superficie agrícola de siembra**”.

GRAFICO LIII: DEL FACTOR I CON EL FACTOR II



Como podemos ver en el GRAFICO LIII, la variable X1 que representa a siembra no esta correlacionada con el resto de variables, encontrándose dispersa de las demás variables.

CAPITULO IV

CONCLUSIONES Y RECOMENDACIONES

4.1 Conclusiones

Al finalizar el estudio en el periodo de referencia de los a los 1989 hasta el 2002 se llegaron a la siguientes conclusiones.

1. Las variables superficie sembrada, superficie cosechada difieren muy poco en hectáreas. Desde el año 1989 hasta el 2002, la menor superpie sembrada tuvo lugar en el año 1994 y la mayor en el año 1997. La superficie cosechada fue mayor en el año 1999 y fue menor en el año 1997. Ambas mantienen un comportamiento inestable, esto es ni creciente ni decreciente, pero se vieron afectadas en el 94 y 97 por los cambios climáticos, provocados por el fenómeno del niño.
2. La mayor superficie sembrada se dio en el año 1997 la cual coincide con el año que fue afectado por el fenómeno del niño, y la menor superficie cosechada fue en este mismo año al igual

que su menor producción, de igual manera se tuvo un mayor roza de cantero por su abundante monte en los sombríos.

3. La menor producción de caña de azúcar tuvo lugar en el año 1997 y la mayor tuvo lugar en el año 1998. La producción tan baja en el año 1997 se debió a los efectos devastadores causados por el fenómeno del niño, ya que el clima produjo la propagación de monte o paja. Sin embargo los cañicultores lograron recuperarse en el año 1998. La cosecha tiene un promedio de 709029.85 Tm por año, pero la mas baja en el periodo de referencia también fue en el año 1997.
4. El menor riego tuvo en el año 1997 debido a que no se tubo la necesidad de regar las hectáreas sembradas por que en ese año fue donde los sombríos estuvieron inundados por efectos del fenómeno de la corriente del niño.
5. El mayor aporque se dio en el año 2001, y el menor herbicida se también en este mismo año lo que podemos deducir que a mayor aporque cuya labor es acumular tierra alrededor de la planta tendríamos un menor fertilizado con herbicida debido a que la planta esta mas resistente a cualquier plaga.
6. El menor escarificado se dio en el año 1997 donde de igual manera en el mismo año se dio el menor riego, sacada de paja

cauca, fertilización, rastrillo y herbicida y en el mismo año se dio la mayor roza de cantero.

7. Por medio del análisis por componentes principales, permitió que de las 11 variables originales nos quedamos con 2 variables ficticias denominadas factores, los cuales explican un 99.96% de la variación total del conjunto de datos, el primer componente muestra mayor ponderación en las variables producción y riego mientras que el segundo tiene un mayor peso en las variables riego y roza.
8. El mismo estudio con datos estandarizados permitió reducir de 12 variables originales a dos factores, estos explicaban el 94.2% de la variabilidad total de los datos. El primer componente representa el eje de producción cosechada, el segundo componente el eje de labores de insumo.

Se aplicó el método de **VARIMAX**, que consiste en minimizar el número de variables que tienen alta carga en un factor con lo cual simplificaríamos la información de la siguiente manera.

El primer factor corresponde a la superficie agrícola de cosecha, ya que tiene una mayor correlación con la variable producción, superficie cosechada, escarificado, roza de cantero.

El segundo factor tiene mayor correlación con la variable superficie sembrada con un coeficiente de 0.99 y un contraste con roza de cantero con un coeficiente de -0.010 por lo cual este factor lo llamaremos superficie agrícola de siembra.

9. La variable X1 que representa a siembra no esta correlacionada con el resto de variables, encontrándose dispersa de las demás variables.

4.2 Recomendaciones

1. Los ingenios deberían realizar encuestas por muestreo en el sector con el objeto de hacer estudios más periódicos de la producción y cultivo de caña de azúcar, y tener información estadística oportuna y confiable.
2. Capacitar mejor a los Ingenieros de campo en el empleo de técnicas estadísticas, estas pueden ser analizadas mediante muestras para así lograr obtener una mejor cosecha y producción.

3. Obtener datos históricos en cuanto a costos de labores de campo para poder realizar un análisis en cuanto a pérdida o ganancia y a su vez reducir la aplicación de fertilizantes y otras labores como rastrillo, herbicidas, roza, escarificado, y utilizarlo mediante una técnica propuesta por medio de este análisis.

ANEXO

DATOS ORIGINALES								
		SUPERFICIE	SUPERFICI		ind-			
	año	SEMBRADA	COSECHADA	PRODUCCION	escarificad	ESCARIFICADO	ind-riego	RIEGO
1	1989	10378	8080,16	567503	0,62	6434,36	2,4	24907,2
2	1990	10380	8063,81	704116	0,6	6228	2,47	25638,6
3	1991	10489	8656,13	718215	0,59	6188,51	2,5	26222,5
4	1992	10490	7763,91	742747	0,72	7552,8	2,57	26959,3
5	1993	8884,49	8473,44	672555	0,72	5953	2,5	20558
6	1994	8221,77	9006,9	667449,56	0,63	6317	2,44	24595
7	1995	10062,44	9231,32	780273,94	0,7	7235	2,89	29965
8	1996	10385,64	10076,44	832253,72	0,77	8607	2,82	31369
9	1997	11114,23	3548,77	405631,13	0,13	1448,65	0,06	621,89
10	1998	11002,71	9561,44	865019,75	0,77	8465,38	2,17	23828,24
11	1999	11002,71	10570,51	700825,72	0,67	7376,07	2,13	23549,01
12	2000	11058,88	9884,62	749487,32	0,75	7859	2,48	26006
13	2001	10500	10179,03	743474,87	0,79	8268	2,58	27136
14	2002	10500	9769,17	776865,9	0,79	7395	2,93	30718

DATOS ORIGINALES											
ind-sacada	SACADA DE PAJA CAUCA	ind-aporque	APORQUE	ind-fertilizaci3n	FERTILIZACI3N	ind-roza de	ROZA DE CANTERO	ind-rastrillo	RASTRILLO	ind-herbicida	HERBICIDA
0,75	7783,5	0,38	3943,64	0,87	9028,86	1,96	20340,88	0,59	6123,02	1,56	16189,68
0,77	7992,6	0,3	3114	0,86	8926,8	1,98	20552,4	0,59	6124,2	1,56	16192,8
0,63	6608,07	0,24	2517,36	0,99	10384,11	2,01	21082,89	0,6	6293,4	1,3	13635,7
0,62	5464	0,09	944,1	0,89	9336,1	2,12	22238,8	0,61	6398,9	1,3	13637
0,81	6686	0,27	2226	0,99	8143	1,74	15431	0,9	7369	1,02	8353
0,81	8101	0,41	4119	0,88	8883	2,03	16726	0,61	6142	1,01	10147
0,91	9479	0,28	2958	0,79	8236	1,83	18444	0,75	7764	1,39	14419
1,24	13769	0,25	2828	1	11150	1,98	20615	0,79	8770	1,57	17475
0,15	1613	0,03	370	0,09	936	2,02	22424,18	0,25	2769	0,64	6992,94
0,67	7410,66	0,06	669	0,75	8293,79	1,16	12722,5	0,78	8607,03	1,02	11250,08
0,72	7966,56	0	0	0,7	7722,07	2,01	22114,62	0,7	7791,49	1,21	13386,24
0,78	8187	0,4	4218,74	0,74	7809	0,03	345,2	0,77	8133,49	0,87	13124
0,85	8947	0,48	5074	0,71	7413	2,08	21806	0,83	8715	0,24	3531
0,83	8701	0,24	2555	0,74	7793	2,1	22086	0,82	8582	1	14956

MATRIZ DE COVARIANZA											
	APORQUE	COSECHA	ESCARIF	FERTILIZ	HERBICID	PRODUCCI	RASTRILL	RIEGO	ROZA	SACA_P	SIEMBRA
APORQUE	2430557,485	2273785,899	2283067,55	2819563,416	5893667,731	159520021,2	2248085	8828573,4	7521120	3564585,614	250414,1303
COSECHA	2273785,899	3032188,792	3084116,21	3967671,087	6346266,235	191739358,9	2778235	12745931	10128929	4057962,428	280681,7115
ESCARIF	2283067,553	3084116,212	3167564,18	4062856,557	6403973,547	194876768,5	2843935	13055331	10220660	4117382,381	326449,2454
FERTILIZ	2819563,416	3967671,087	4062856,56	5411613,547	8005045,353	252028515	3549256	17119640	13052246	5520565,719	263603,1851
HERBICID	5893667,731	6346266,235	6403973,55	8005045,353	15410092,38	429986600,3	6027170	25737594	21653066	9240638,117	504221,4399
PRODUCCI	159520021,2	191739358,9	194876769	252028515	429986600,3	12959255537	1,77E+08	806898584	6,46E+08	278160484	14369852,57
RASTRILL	2248084,684	2778235,432	2843934,54	3549256,238	6027170,111	176597517,2	2665067	11321817	9097018	3704991,498	456704,9553
RIEGO	8828573,364	12745931,1	13055331,2	17119640,46	25737594,01	806898584,3	11321817	55570286	42961998	16983261,34	627546,0654
ROZA	7521119,623	10128929,25	10220659,8	13052246,09	21653065,72	645850006,4	9097018	42961998	35401314	13173606	339492,0514
SACA_P	3564585,614	4057962,428	4117382,38	5520565,719	9240638,117	278160484	3704991	16983261	13173606	6724059,394	231970,5011
SIEMBR	250414,1303	280681,7115	326449,245	263603,1851	504221,4399	14369852,57	456705	627546,07	339492,1	231970,5011	674773,2586

MATRIZ DE CORRELACIÓN											
	APORQUE	COSECHA	ESCARIF	FERTILIZ	HERBICID	PRODUCCI	RASTRILL	RIEGO	ROZA	SACA_P	SIEMBRA
APORQUE	1	0,837564962	0,82281671	0,777437582	0,963008406	0,898818957	0,883295	0,7596548	0,810811	0,88173957	0,195536101
COSECHA	0,837564962	1	0,99515311	0,979476805	0,928405164	0,96725977	0,97732	0,981911	0,977633	0,898698739	0,196226232
ESCARIF	0,822816707	0,995153107	1	0,981308114	0,916609167	0,961849959	0,97882	0,9840199	0,965176	0,892159942	0,223292443
FERTILIZ	0,777437582	0,979476805	0,98130811	1	0,876592576	0,951691666	0,934587	0,9872109	0,943001	0,915175586	0,137945753
HERBICID	0,963008406	0,928405164	0,91660917	0,876592576	1	0,962193179	0,940496	0,8795165	0,927058	0,907786189	0,156365074
PRODUCCI	0,898818957	0,96725977	0,96184996	0,951691666	0,962193179	1	0,950256	0,9508404	0,953525	0,942300965	0,15366797
RASTRILL	0,883294757	0,977320168	0,97881996	0,934587356	0,940495588	0,950255739	1	0,930338	0,936559	0,875220295	0,340567128
RIEGO	0,759654766	0,981911017	0,98401987	0,98721089	0,879516538	0,950840369	0,930338	1	0,96862	0,878585452	0,102481474
ROZA	0,810811425	0,977632724	0,96517605	0,943001053	0,927058382	0,953524713	0,936559	0,9686201	1	0,853845338	0,069460994
SACA_P	0,88173957	0,898698739	0,89215994	0,915175586	0,907786189	0,942300965	0,87522	0,8785855	0,853845	1	0,108902552
SIEMBR	0,195536101	0,196226232	0,22329244	0,137945753	0,156365074	0,15366797	0,340567	0,1024815	0,069461	0,108902552	1

BIBLIOGRAFIA

1. SICA.GOV.EC. Página web del Ministerio de Agricultura del Ecuador.
2. Jonson A Richard & Wichern W Dean. Applied Multivariate Statistical Analysis. J Prentice Hall, Cuarta Edición 1998.
3. Hair- Anderson- Taham Black. Multivariate Data Analysis. Prentice Hall, 1998.
4. Spigel Murray E. Estadística de Shaum. Mc Graw Hill. Interamericana de México. Segunda Edición 1991.
5. Pla. Laura. Análisis Multivariado. Método de componentes principales, 1986.
6. Mendenhall William, Wackerly Denos D, Scheaffer Richard Estadística Matemática con Aplicaciones Grupo Editorial Iberoamericana S.A, Segunda Edición 1984.