

## USO DE LOS MODULOS DE SPSS

<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
<b>1. INFORMES :</b> <ul style="list-style-type: none"> <li>• Cubos OLAP</li> <li>• Resúmenes de casos</li> <li>• Informe de estadísticos en fila</li> <li>• Informe de estadísticos en columnas</li> </ul>	<b>BASE</b> Con estos procedimientos se realizan Reportes con la inclusión de variables deseadas. Este es el REPORTEADOR de SPSS. Con los cubos OLAP se hacen cientos de tablas con todos los cruces deseados. Es muy útil para concentrar toda la información con variables Categóricas y Numéricas.
<b>2. ESTADISTICOS DESCRIPTIVOS</b> <ul style="list-style-type: none"> <li>• Frecuencias</li> <li>• Descriptivos</li> <li>• Explorar</li> <li>• Tablas de Contingencia</li> </ul>	<b>BASE</b> Aquí se encuentran las herramientas para analizar una por una de todas las variables. La tabla de contingencia sirve para comprobar que las variables de Fila y Columna son independientes o se relacionan.
<b>3. TABLAS</b> <ul style="list-style-type: none"> <li>• Tablas Personalizadas</li> <li>• Conjuntos de Respuestas Múltiples</li> </ul>	<b>TABLAS</b> Este módulo sirve para hacer los cruces de variables de una forma rápida y fácil. Es la herramienta ideal para presentar informes que contengan muchas tablas.
<b>4. COMPARAR MEDIAS :</b> <ul style="list-style-type: none"> <li>• Medias</li> <li>• Prueba T para 1 muestra</li> <li>• Prueba T para muestras Independientes</li> <li>• Prueba T para muestras Relacionadas</li> <li>• ANOVA de un factor</li> </ul>	<b>BASE</b> Con estos procedimientos se analiza si 2 productos o eventos son suficientemente diferentes (estadísticamente hablando) para afirmar que son percibidos en forma diferente. Se realizan pruebas para investigar en que difieren. ANOVA de un factor genera un análisis de varianza para una variable dependiente cuantitativa respecto a una única variable o factor (la variable independiente). Se utiliza para contrastar la hipótesis de que varias medias son iguales. Esta técnica es una extensión de la prueba t para dos muestras.
<b>5. MODELO LINEAL GENERAL :</b> <ul style="list-style-type: none"> <li>• Univariante</li> </ul> <hr style="border-top: 1px dashed black;"/> <ul style="list-style-type: none"> <li>• Multivariante</li> <li>• Medidas Repetidas</li> <li>• Componentes de la Varianza</li> </ul>	<b>BASE</b> MLG Univariante proporciona un análisis de regresión y un análisis de varianza para una variable dependiente de escala mediante uno o más factores o variables. Util para hallar la variación de un fenómeno por un factor.
	<b>MODELOS AVANZADOS</b> En cambio el MLG Multivariante se utiliza con variables dependientes múltiples. Se utiliza para predecir un evento con base en varios factores.
<b>6. CORRELACIONES :</b> <ul style="list-style-type: none"> <li>• Bivariadas</li> <li>• Parciales</li> <li>• Distancias</li> </ul>	<b>BASE</b> Es útil para medir cómo están relacionadas 2 variables. Si la relación es fuerte y si es una relación positiva o negativa (si aumenta una aumenta la otra o la inversa)

<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
<b>7. REGRESION :</b> <ul style="list-style-type: none"> <li>Lineal</li> <li>Estimación Curvilínea</li> </ul>	<p style="text-align: center;"><b>BASE</b></p> <p>La regresión Lineal sirve para pronosticar valores entre la relación de varias variables numéricas</p>
<ul style="list-style-type: none"> <li>Logística binaria</li> <li>Logística multinomial</li> <li>Ordinal</li> <li>Probit</li> <li>No lineal</li> <li>Estimación Ponderada</li> <li>Mínimos Cuadrados en 2 fases</li> <li>Escalamiento Optimo</li> </ul>	<p style="text-align: center;"><b>MODELOS DE REGRESION</b></p> <p>La regresión Logística sirve para pronosticar si un evento se produce o no (la salida es una variable dicotómica).</p> <p><b>PROBIT</b> mide la relación entre la intensidad de un estímulo y la proporción de casos que presentan una cierta respuesta a dicho estímulo. Es útil para las situaciones en las que se dispone de una respuesta dicotómica que se piensa puede estar influenciada o causada por los niveles de alguna o algunas variables independientes.</p> <p><b>Escalamiento Optimo</b> es para los datos categóricos que se utilizan con frecuencia en los estudios de mercado, los estudios de encuestas y la investigación en las ciencias sociales y del comportamiento con conjuntos de datos con las siguientes características:</p> <ul style="list-style-type: none"> <li>Observaciones insuficientes</li> <li>Demasiadas variables</li> <li>Demasiados valores por cada variable</li> </ul>
<b>8. LOGLINEAL :</b> <ul style="list-style-type: none"> <li>General</li> <li>Logit</li> <li>Selección de Modelo</li> </ul>	<p style="text-align: center;"><b>MODELOS AVANZADOS</b></p> <p>El Análisis loglineal general analiza las frecuencias de las observaciones incluidas en cada categoría de la clasificación cruzada de una tabla de contingencia. Identifica los factores explicativos de relaciones de casos.</p> <p><b>Logit</b> analiza la relación entre variables dependientes Categóricas (o de respuesta) y varias variables independientes (o explicativas).</p>
<b>9. CLASIFICAR :</b> <ul style="list-style-type: none"> <li>Conglomerado de K medias</li> <li>Conglomerados Jerárquicos</li> <li>Discriminante</li> </ul>	<p style="text-align: center;"><b>BASE</b></p> <p><b>Conglomerados de K Medias</b> intenta identificar grupos de casos relativamente homogéneos basándose en las características o variables seleccionadas. Puede gestionar un gran número de casos. Muy útil para descubrir grupos no fáciles de ver dentro de una base de datos.</p> <p>El Análisis de Conglomerados Jerárquico descubre conglomerados en pocos casos (menos de 100)</p> <p>El Análisis Discriminante resulta útil para construir un modelo predictivo para pronosticar el grupo de pertenencia de un caso a partir de las características observadas de cada caso.</p>
<b>10. REDUCCION DE DIMENSIONES:</b> <ul style="list-style-type: none"> <li>Análisis Factorial</li> </ul>	<p style="text-align: center;"><b>BASE</b></p> <p>El análisis Factorial agrupa muchas variables en Factores o variables subyacentes de temas comunes. Muy útil para analizar por temas un gran No. De variables respecto a un fenómeno observado.</p>
<ul style="list-style-type: none"> <li>Análisis de Correspondencias</li> <li>Escalamiento Optimo</li> </ul>	<p style="text-align: center;"><b>CATEGORIAS</b></p> <p>El Análisis de Correspondencias (ANACOR) describe las relaciones existentes entre variables nominales sobre un espacio de 2</p>

<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
<ul style="list-style-type: none"> <li>• HOMALS</li> <li>• CATPCA</li> <li>• OVERALS</li> </ul>	<p>dimensiones llamado Mapa Perceptual.</p> <p>El escalamiento multidimensional encuentra la estructura existente en un conjunto de medidas de proximidades entre objetos. Asigna las observaciones a posiciones específicas en un espacio conceptual de pocas dimensiones (Mapa) , de modo que las distancias entre los puntos en el espacio concuerden al máximo con las similitudes (o disimilitudes) dadas.</p>
<p><b>11. ESCALAS :</b></p> <ul style="list-style-type: none"> <li>• Análisis de Fiabilidad</li> <li>• Escalamiento Multidimensional</li> </ul>	<p style="text-align: center;"><b>BASE</b></p> <p>El análisis de Fiabilidad permite estudiar las propiedades de las escalas de medición (Escala de Likert) y de los elementos que componen las escalas. Mide la confiabilidad con que se responden este tipo de preguntas en una encuesta.</p> <p>El Escalamiento Multidimensional encuentra la estructura de un conjunto de medidas de distancia entre objetos o casos en un espacio conceptual (normalmente de dos o tres dimensiones). Trata datos de disimilitud procedentes de múltiples fuentes, como podrían ser múltiples evaluadores o múltiples sujetos evaluados por un cuestionario.</p>
<ul style="list-style-type: none"> <li>• Escalamiento Multidimensional PROXSCAL</li> </ul>	<p style="text-align: center;"><b>CATEGORIAS</b></p> <p>Encuentra la estructura existente en un conjunto de medidas de proximidades entre objetos. Se asignan las observaciones a posiciones específicas en un espacio conceptual de pocas dimensiones, de modo que las distancias entre los puntos en el espacio concuerden al máximo con las similitudes (o disimilitudes) dadas. El resultado es una representación de mínimos cuadrados de los objetos en dicho espacio que ayudará a entender mejor los datos. Es útil en la determinación de relaciones perceptuales.</p>
<p><b>12. PRUEBAS NO PARAMETRICAS</b></p> <ul style="list-style-type: none"> <li>• Chi-cuadrado</li> <li>• Binomial</li> <li>• Rachas</li> <li>• K-S de 1 muestra</li> <li>• 2 Muestras Independientes</li> <li>• K muestras Independientes</li> <li>• 2 muestras relacionadas</li> <li>• K muestras relacionadas</li> </ul>	<p style="text-align: center;"><b>BASE</b></p> <p>Son pruebas para realizarse en muestras pequeñas (menos de 100 casos) y no requieren supuestos sobre la forma de la distribución subyacente. Se asume que los datos son una muestra aleatoria.</p>
<p><b>13. SERIES TEMPORALES :</b></p> <ul style="list-style-type: none"> <li>• Suavizado Exponencial</li> <li>• Autoregresión</li> <li>• ARIMA</li> <li>• Descomposición Estacional</li> </ul>	<p style="text-align: center;"><b>TENDENCIAS</b></p> <p>El análisis de series temporales intenta predecir los valores futuros de una serie de datos medidos con un intervalo regular (meses, días, horas, etc.) a lo largo de mucho tiempo. El modelo que se genera debe explicar los valores pasados para predecir si aumentarán o disminuirán los próximos valores y en qué medida lo harán. La capacidad de realizar dichas predicciones correctamente es muy importante para cualquier negocio o institución.</p> <p>Por medio del Modelizador de series temporales estima el modelo de suavizado exponencial, el modelo autorregresivo integrado de</p>

<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
	<p>media móvil (ARIMA) univariante y los modelos ARIMA (o modelos de función de transferencia) multivariantes para series temporales, y genera predicciones. Estima el mejor ajuste para una o más series de variables dependientes, lo que elimina la necesidad de identificar un modelo adecuado mediante ensayo y error.</p>
<p><b>14. SUPERVIVENCIA :</b></p> <ul style="list-style-type: none"> <li>• Tablas de Mortalidad</li> <li>• Kaplan-Meier</li> <li>• Regresión de Cox</li> <li>• Cox con covariable dep. del tiempo</li> </ul>	<p style="text-align: center;"><b>MODELOS AVANZADOS</b></p> <p>Proporcionan procedimientos que ofrecen opciones de modelado más avanzadas que las disponibles en el módulo Base.</p> <ul style="list-style-type: none"> <li>• <b>MLG Multivariado</b> amplía el modelo lineal general que proporciona <b>MLG Univariado</b> al permitir varias variables dependientes. Una extensión adicional, <b>GLM Medidas repetidas</b>, permite las medidas repetidas de varias variables dependientes.</li> <li>• <b>Análisis de componentes de la varianza</b> es una herramienta específica para descomponer la variabilidad de una variable dependiente en componentes fijos y aleatorios.</li> <li>• Los <b>modelos mixtos lineales</b> amplían el modelo lineal general de manera que los datos puedan presentar variabilidad corelacionada y no constante. El modelo lineal mixto proporciona, por tanto, la flexibilidad necesaria para modelar no sólo las medias sino también las varianzas y covarianzas de los datos.</li> <li>• Los <b>modelos lineales generalizados (GZLM)</b> relajan el supuesto de normalidad del término de error y sólo requieren que la variable dependiente esté relacionada linealmente con los predictores mediante una transformación o función de enlace. Las ecuaciones de estimación generalizada (<b>GEE</b>) amplía <b>GZLM</b> para permitir medidas repetidas.</li> <li>• El <b>análisis loglineal general</b> permite ajustar modelos a datos de recuento de clasificación cruzada y la selección del modelo del análisis loglineal puede ayudar a elegir entre modelos.</li> <li>• El <b>análisis loglineal LOGIT</b> permite ajustar modelos loglineales para analizar la relación existente entre una variable dependiente categórica y uno o más predictores categóricos.</li> <li>• Se puede realizar un análisis de supervivencia a través de <b>Tablas de mortalidad</b> para examinar la distribución de variables de tiempo de espera hasta un evento, posiblemente por niveles de una variable de factor. La <b>regresión de Cox</b> se utiliza para modelar el tiempo de espera hasta un determinado evento, basado en los valores de las covariables especificadas.</li> </ul>
<p><b>15. RESPUESTAS MÚLTIPLES :</b></p> <ul style="list-style-type: none"> <li>• Definir Conjuntos</li> <li>• Frecuencias</li> <li>• Tablas de Contingencia</li> </ul>	<p style="text-align: center;"><b>BASE</b></p> <p>Los conjuntos de respuestas múltiples utilizan varias variables para registrar respuestas a preguntas en las que el encuestado puede ofrecer más de una respuesta. Se definen los conjuntos y son nuevas variables "Virtuales" sin que se vean en el Visor de datos. Los conjuntos de respuestas múltiples se consideran variables categóricas y se les pueden aplicar los siguientes procedimientos generales para este tipo de variable : <b>Frecuencias, Tablas de</b></p>

<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
	Contingencias.
<p><b>16. TABLAS PERSONALIZADAS :</b></p> <ul style="list-style-type: none"> <li>• Tablas simples para variables Categóricas</li> <li>• Tablas complejas de variables categóricas y De Escala</li> <li>• Tablas con datos en Filas, Columnas y Capas (3 dimensiones de tablas Pivotales)</li> <li>• Estadísticos intuitivos de acuerdo al nivel de medición de la variable</li> <li>• Respuestas múltiples</li> <li>• Estadísticos de contraste</li> </ul>	<p style="text-align: center;"><b>TABLAS</b></p> <p>Es un módulo muy útil para producir todas las salidas de cruces de variables que se necesitan hacer en forma intuitiva (se previsualizan las tablas) y flexible.</p> <p>Se encuentran disponibles tres pruebas de significación distintas, para poder estudiar las relaciones existentes entre las variables de filas y de columnas dependiendo de los efectos de la anidación y la apilación de las variables.</p> <p>Las Respuestas Múltiples son preguntas con un conjunto común de respuestas posibles que se analizan en forma especial y fácil con este módulo Tablas.</p> <p>Las capas de Tablas añaden una dimensión de profundidad a las tablas, creando "cubos" tridimensionales. Las capas son similares a la anidación o la apilación. La diferencia fundamental es que sólo resulta visible simultáneamente una categoría de capa.</p> <p>Los estadísticos básicos disponibles para las variables categóricas son las frecuencias y los porcentajes. También se pueden especificar estadísticos de resumen personalizados para los totales y subtotales. Estos estadísticos de resumen personalizados incluyen medidas de tendencia central (como la media y la mediana) y dispersión (como la desviación típica) que pueden ser adecuadas para algunas variables categóricas ordinales.</p>
<p><b>17. ARBOLES DE CLASIFICACION</b></p> <ul style="list-style-type: none"> <li>• Método CHAID</li> <li>• Método CHAID Exhaustivo</li> <li>• CRT</li> <li>• QUEST</li> </ul>	<p style="text-align: center;"><b>ARBOLES DE CLASIFICACION</b></p> <p>Crean un modelo de clasificación basado en árboles, y clasifica casos en grupos o pronostica valores de una variable (criterio) dependiente basada en valores de variables independientes (predictores). El procedimiento proporciona herramientas de validación para análisis de clasificación exploratorios y confirmatorios.</p> <p>Se utiliza en Segmentación y es un módulo muy útil para analizar grandes cantidades de variables (50, 100) y descubrir cuáles son las más importantes, cuáles en segundo lugar y así sucesivamente en función de la significancia detectada por el estadístico Chi Cuadrado.</p> <p>Los métodos de crecimiento del árbol ( o Modelos) disponibles son: <b>CHAID</b>. Detección automática de interacciones mediante chi-cuadrado (Chi-square Automatic Interaction Detection). En cada paso, CHAID elige la variable independiente (predictora) que presenta la interacción más fuerte con la variable dependiente. Las categorías de cada predictor se funden si no son significativamente distintas respecto a la variable dependiente.</p> <p><b>CHAID exhaustivo</b>. Una modificación del CHAID que examina todas las divisiones posibles de cada predictor.</p> <p><b>CRT</b>. Árboles de clasificación y regresión (Classification and Regression Trees). CRT divide los datos en segmentos para que sean lo más homogéneos que sea posible respecto a la variable</p>

<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
	<p>dependiente. Un nodo terminal en el que todos los casos toman el mismo valor en la variable dependiente es un nodo homogéneo y "puro".</p> <p><b>QUEST.</b> Árbol estadístico rápido, insesgado y eficiente (Quick, Unbiased, Efficient Statistical Tree). Un método que es rápido y que evita el sesgo que presentan otros métodos al favorecer los predictores con muchas categorías. Sólo puede especificarse QUEST si la variable dependiente es nominal.</p> <p>En ocasiones, los modelos de árbol pueden contener tantos nodos y ramas que resulta difícil o imposible ver todo el árbol a tamaño completo. Para ello existen ciertas funciones que son de utilidad a la hora de trabajar con árboles grandes:</p> <ul style="list-style-type: none"> <li>• <b>Mapa del árbol.</b> Puede utilizar el mapa del árbol, que es una versión más pequeña y simplificada del árbol, para desplazarse por él y seleccionar nodos.</li> <li>• <b>Escalamiento.</b> Puede acercarse o alejarse cambiando el porcentaje de escala para la presentación del árbol.</li> <li>• <b>Presentación de nodos y ramas.</b> Se puede hacer que la presentación de un árbol sea más compacta mostrando sólo tablas o sólo gráficos en los nodos, o desactivando la visualización de las etiquetas de los nodos o la información de las variables independientes.</li> <li>• <b>Reglas de selección y puntuación :</b> Ofrece la capacidad de generar reglas de selección o clasificación/predicción en forma de sintaxis de comandos de SPSS, SQL o sólo texto (inglés sin formato). Estas reglas se pueden visualizar en el Visor y/o guardar en un archivo externo.</li> </ul>
<p><b>18. CONJOINT (CoIntegración)</b></p> <ul style="list-style-type: none"> <li>• Generador de diseños (Orthoplan)</li> <li>• Plancards (generador de tarjetas)</li> <li>• Análisis Conjoint</li> </ul>	<p style="text-align: center;"><b>CONJOINT</b></p> <p>SPSS Conjoint o Análisis Conjunto es un módulo ideal para optimizar las opciones o alternativas de escogencia de consumidores en diseño de nuevos productos o en la modificación de los actuales.</p> <p>El análisis conjoint permite conocer con detalle las preferencias de los consumidores, sus intenciones de compra, así como para simular cómo reaccionarían a los cambios en los actuales productos o servicios o a la introducción de nuevos servicios en un entorno competitivo.</p> <p>En un estudio de conjoint a los participantes se les muestran varias combinaciones de opciones y atributos (generados al azar) para que las evalúen de forma simultánea. Mediante un tamaño de muestra relativamente pequeño el investigador de mercados puede obtener información muy útil sobre el valor relativo que asignan los consumidores a los diferentes atributos ofrecidos por las distintas opciones presentadas.</p> <p>El generador de diseños (Orthoplan) genera un vector ortogonal de posibles productos alternativos que combinan diferentes características de los productos con los niveles especificados.</p> <p>Plancards, genera tarjetas que los encuestados pueden ordenar con</p>

<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
	<p>el fin de clasificar productos alternativos.</p> <p>Analice los datos con el procedimiento Conjoint, una versión especialmente adaptada de la regresión. Obtendrá resultados con los que podrá actuar, como por ejemplo qué características de un producto son importantes y qué niveles son preferidos en ellas. También podrá hacer simulaciones que le indicarán la cuota de mercado de los productos alternativos.</p> <p>4 puntos que harán el lanzamiento de un producto todo un éxito</p> <ol style="list-style-type: none"> <li>1. El producto correcto – Diseñe el producto con el conjunto de características que el mercado busca</li> <li>2. El precio correcto – Establezca el precio basándose en el valor que el mercado le ha asignado</li> <li>3. El lugar correcto – Prediga cómo la combinación precio – producto se desempeñará en el mercado antes de establecer los recursos para lanzar el producto</li> <li>4. La promoción correcta – Enfoque sus esfuerzos de marketing en los atributos o características que más le interesan a sus clientes.</li> </ol>
<p><b>19. MUESTRAS COMPLEJAS</b></p> <ul style="list-style-type: none"> <li>• Estratificación</li> <li>• Conglomerados</li> <li>• Múltiples etapas</li> </ul>	<p style="text-align: center;"><b>MUESTRAS COMPLEJAS</b></p> <p>Muestras complejas de SPSS permite seleccionar una muestra de acuerdo con un diseño complejo e incorporar las especificaciones del diseño al análisis de los datos, asegurando así que los resultados serán válidos.</p> <p>Una muestra compleja determinada puede tener alguna o todas las características siguientes:</p> <p><b>Estratificación.</b> El muestreo estratificado implica seleccionar muestras independientemente dentro de los subgrupos de la población que no se solapan o estratos.</p> <p><b>Conglomerados.</b> El muestreo por conglomerados implica la selección de grupos de unidades muestrales o conglomerados.</p> <p><b>Múltiples etapas.</b> En el muestreo polietápico, se selecciona una muestra de primera etapa basada en conglomerados. A continuación, se crea una muestra de segunda etapa extrayendo submuestras a partir de los conglomerados seleccionados.</p> <p><b>Plan de muestreo.</b> Las especificaciones dadas en el Asistente de muestreo definen un diseño muestral que se utiliza para extraer una muestra compleja. El archivo del plan de muestreo contiene esas especificaciones. El archivo del plan de muestreo también contiene un plan de análisis por defecto que utiliza métodos de estimación adecuados para el diseño muestral especificado.</p> <p><b>Plan de análisis.</b> Este archivo de plan contiene la información necesaria en los procedimientos de análisis de Muestras complejas para calcular correctamente las estimaciones de la varianza de una muestra compleja. El plan incluye la estructura de la muestra, los métodos de estimación de cada etapa y las referencias para variables necesarias como por ejemplo, las ponderaciones muestrales. El Asistente de preparación del análisis permite crear y editar los planes de análisis.</p>

<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
<p><b>20. PRUEBAS EXACTAS</b></p> <ul style="list-style-type: none"> <li>• Método Exacto</li> <li>• Prueba de Monte Carlo</li> </ul>	<p style="text-align: center;"><b>PRUEBAS EXACTAS</b></p> <p>Pruebas exactas proporciona dos métodos adicionales para calcular los niveles de significación de los estadísticos disponibles mediante los procedimientos Tablas de contingencia y Pruebas no paramétricas. Estos métodos, el método exacto y el de Monte Carlo, proporcionan el medio para obtener resultados exactos cuando los datos no cumplen alguno de los supuestos subyacentes necesarios para obtener resultados fiables con el empleo del método asintótico típico.</p> <p>Los resultados de Monte Carlo pueden ser obtenidos con mayor rapidez que los resultados exactos.</p> <p>En general, las pruebas exactas se pueden realizar rápidamente con tamaños muestrales por debajo de 30 casos.</p>
<p><b>21. VALORES PERDIDOS</b></p> <ul style="list-style-type: none"> <li>• % de discordancia</li> <li>• Pruebas "t"</li> <li>• Tablas de Contingencia</li> </ul>	<p style="text-align: center;"><b>VALORES PERDIDOS</b></p> <p>El módulo Análisis de Valores Perdidos realiza tres funciones principales:</p> <ul style="list-style-type: none"> <li>• Describe el patrón de los datos perdidos: la ubicación de los mismo, cuán extensos son, si los pares de variables tienden a tener valores perdidos en casos diferentes, si los valores de los datos son extremos y si los valores están perdidos de forma aleatoria.</li> <li>• Estima medias, desviación típica, covarianzas y correlaciones utilizando un método por lista, por parejas, de regresión, o bien EM (expectation-maximization, maximización esperada). El método por parejas muestra, además, recuentos de los casos completos por parejas.</li> <li>• Rellena (imputa) los valores perdidos con valores utilizando el método EM o el de regresión.</li> </ul> <p>Este análisis ayuda a resolver varios problemas ocasionados por los datos incompletos. Los casos con valores perdidos que son sistemáticamente diferentes de los casos sin valores perdidos dificultan la interpretación de los resultados.</p> <p>Además, los datos perdidos pueden reducir la precisión de los estadísticos calculados, porque no se dispone de tanta información como originalmente se pensaba. Otro problema radica en que los supuestos subyacentes a muchos procedimientos estadísticos se basan en casos completos y los valores perdidos pueden complicar la teoría exigida.</p> <p>Para cada variable, SPSS crea una variable de indicador de valores perdidos, que indica si el valor de la variable está presente o está perdido. Las variables de indicador no se muestran, pero se utilizan para crear el porcentaje de discordancia, las pruebas t y las tablas de frecuencia.</p> <ul style="list-style-type: none"> <li>• Porcentaje de discordancia. Para cada par de variables muestra el porcentaje de casos en los que una variable tiene un valor perdido y la otra variable tiene un valor no perdido.</li> </ul>



<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
	<p>Cada elemento diagonal de la tabla contiene el porcentaje de valores perdidos para una sola variable.</p> <ul style="list-style-type: none"> <li>• Pruebas t con los grupos formados por las variables de indicador. Se comparan las medias de los dos grupos para cada variable cuantitativa, utilizando el estadístico t de Student. Los grupos se determinan según si la variable de indicador está codificada como presente o como perdida. Se muestra el estadístico t, los grados de libertad, los recuentos de valores perdidos y no perdidos y las medias de los dos grupos.</li> <li>• Tablas de contingencia de variables categóricas y de indicador. Para cada variable categórica se muestra una tabla. Para cada categoría, la tabla muestra la frecuencia y el porcentaje de los valores no perdidos para las demás variables. También se muestran los porcentajes de cada tipo de valor perdido.</li> </ul>
<p><b>22. PREPARACION DE DATOS</b></p> <ul style="list-style-type: none"> <li>• Preparación de Metadatos</li> <li>• Validación de datos</li> <li>• Preparación de Modelos</li> <li>• Reglas de Validación</li> </ul>	<p style="text-align: center;"><b>PREPARACION DE DATOS</b></p> <p>A medida que la potencia de los sistemas informáticos se incrementa, la necesidad de información crece proporcionalmente, llevando a un crecimiento cada vez mayor de la recopilación de datos: más casos, más variables y más errores en la entrada de datos. Estos errores son la pesadilla de las predicciones del modelo predictivo, que son el objetivo final del almacenamiento de datos, por lo que existe la necesidad de mantener los datos "limpios". Sin embargo, la cantidad de datos almacenados ha superado de tal forma a la capacidad de comprobar los casos manualmente que resulta vital implementar procesos automatizados para validar los datos.</p> <p>El módulo adicional de SPSS Data Preparation permite identificar casos, variables y valores de datos atípicos y no válidos en el conjunto de datos activo.</p> <p>Procedimientos de preparación de datos:</p> <ul style="list-style-type: none"> <li>• <u>Preparación de metadatos.</u> Revisa las variables del archivo de datos y determina los valores válidos, las etiquetas y los niveles de medidas. Identifica las combinaciones de valores de variable que son imposibles pero suelen estar mal codificadas. Define las reglas de validación en función de esta información. Esta tarea puede resultar pesada, pero el esfuerzo compensa si debe validar archivos de datos que tengan atributos similares con regularidad.</li> <li>• <u>Validación de datos.</u> Ejecuta comprobaciones básicas y comprobaciones de reglas de validación definidas para identificar casos no válidos, variables y valores de datos. Cuando se encuentran datos no válidos, investiga y corrige la causa. Puede que sea necesario realizar otro paso con la preparación de metadatos.</li> <li>• <u>Preparación de modelos.</u> Identifica valores atípicos estadísticos potenciales que puedan provocar problemas para muchos modelos predictivos. Algunos valores atípicos son el resultado de valores de variable no válidos que no se</li> </ul>

<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
	<p>han identificado. Puede que sea necesario realizar otro paso con la preparación de metadatos.</p> <p>Se pueden especificar reglas de Validación para determinar si un caso es válido. Existen dos tipos de reglas de validación:</p> <ul style="list-style-type: none"> <li>• <b>Reglas de variable única.</b> Las reglas de variable única constan de un conjunto fijo de comprobaciones que se aplican a una única variable, como las comprobaciones de los valores que están fuera de rango. En el caso de las reglas de variable única, los valores válidos pueden expresarse como un rango de valores o una lista de valores aceptables.</li> <li>• <b>Reglas inter-variables.</b> Las reglas inter-variables son reglas definidas por el usuario que se pueden aplicar a una única variable o a una combinación de variables. Las reglas inter-variables están definidas por una expresión lógica que marca valores no válidos.</li> </ul> <p>Las reglas de validación se guardan en el diccionario de datos del archivo de datos. Esto permite especificar una regla una vez y volver a utilizarla más adelante</p> <p>Una vez que el archivo de datos está "limpio", se pueden generar modelos de otros módulos de SPSS.</p>
<p><b>23. AMOS</b></p> <ul style="list-style-type: none"> <li>• Diagramas de camino</li> <li>• Genera modelos de ecuaciones estructurales (incluyendo análisis causal y modelos de datos longitudinales) con variables observables y latentes</li> <li>• Modelamiento individual como un conjunto de restricciones de los parámetros del modelo</li> <li>• Análisis exploratorio.</li> <li>• Criterio de información de Akaike (AIC)</li> <li>• Criterio de información bayesiano (BIC) para comparar los modelos.</li> <li>• Análisis factoriales confirmatorios</li> <li>• Modelos de los componentes de la varianza</li> <li>• Modelos de errores de las variables</li> <li>• Modelos de variable general latente.</li> <li>• Estructuras promedio y conjuntos de datos multigrupo</li> <li>• Especificación automática</li> <li>• Análisis simultáneo de</li> </ul>	<p style="text-align: center;"><b>AMOS</b> (Ecuaciones Estructurales)</p> <p>Amos es un software de modelización de ecuaciones estructurales (Structural Equation Modeling – SEM) que permite dar apoyo a investigaciones y teorías al ir más allá de los métodos de análisis multivariante, regresión, análisis factorial, correlación o de análisis de la varianza. Con Amos se puede especificar, estimar, evaluar y presentar un modelo en un diagrama intuitivo, que mostrará las relaciones posibles entre las variables, confirmando relaciones complejas.</p> <p>Con Amos se pueden desarrollar modelos de actitud y de comportamiento que reflejen relaciones complejas de forma realista. Cualquier variable numérica, tanto observada como latente, se podrá utilizar para predecir cualquier otra variable numérica. Amos incluye opciones ampliadas de estadística basadas en la estimación Bayesiana, se pueden llevar a cabo estimaciones con datos categóricos ordinales y datos censurados. Esto permite crear modelos basados en datos no numéricos sin tener que asignar puntuaciones numéricas a los datos, y trabajar con datos censurados sin tener que hacer más asunciones que la de normalidad.</p> <p>Se pueden estimar distribuciones a posteriori predictivas para determinar valores probables para datos perdidos o parcialmente perdidos dentro de un modelo de variable latente.</p> <p>Amos permite analizar datos de distintas poblaciones simultáneamente, tales como grupos étnicos múltiples.</p> <p>Se puede utilizar la técnica de exploración de Amos, la búsqueda de</p>

<b>Procedimiento Estadístico</b>	<b>Módulo de SPSS</b>
<p>datos procedentes de varias poblaciones</p> <ul style="list-style-type: none"> <li>• Combinación de modelos factoriales y de regresión en un único modelo y ajustes de forma simultánea</li> <li>• Modelos Anidados</li> </ul>	<p>especificaciones de modelos de ecuaciones estructurales (SEM), para elegir un modelo entre un amplio abanico de opciones.</p> <p>Aproveche las búsquedas anteriores especificando limitaciones a los valores de parámetros de su modelo, o use la estimación bayesiana para especificar una distribución informativa a priori de los parámetros. Utilice el análisis factorial confirmatorio de Amos para especificar y probar un patrón factorial en vez de tener que recurrir al análisis factorial exploratorio tradicional.</p> <p>También puede incluir varios modelos en un único análisis. Amos examina todos los pares de modelos en los que se puede obtener un modelo aplicando restricciones de parámetros al otro. Amos incluso sugiere cómo se podría mejorar el modelo- por ejemplo, añadiendo una flecha para conectar dos variables. Los gráficos y las estadísticas ayudan a encontrar el equilibrio óptimo entre la sencillez del modelo y su idoneidad.</p> <p>Descubra relaciones imprevistas usando diagramas de camino : después de aceptar el modelo adecuado, el diagrama de camino de Amos muestra la fortaleza de la relación entre variables. Por ejemplo, si está trabajando con datos de una encuesta sobre condimentos, puede suponer inicialmente que la variable "satisfacción del sabor" es el mejor indicador de fidelidad a la marca. Sin embargo, observando el diagrama de camino de Amos, verá que el mejor indicador de fidelidad a la marca es "tamaño del paquete comprado".</p> <p>Amos es el complemento ideal para los análisis en que se utiliza SPSS Modelos Avanzados o SPSS Modelos de Regresión, ya que los modelos de ecuaciones estructurales (SEM) amplían los algoritmos disponibles en esos módulos.</p>