

COMPARACIÓN DE METODOLOGÍAS PARA LA EVALUACIÓN DE EFECTOS EN
LA SATISFACCIÓN GENERADOS POR LAS VARIABLES TEMPO Y MODO
MUSICAL

SANTIAGO ALBERTO REYES GONZÁLEZ

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
FACULTAD DE CIENCIAS BÁSICAS
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA
BOGOTÁ D.C.
2017

COMPARACIÓN DE METODOLOGÍAS PARA LA EVALUACIÓN DE EFECTOS EN
LA SATISFACCION GENERADOS POR LAS VARIABLES TEMPO Y MODO
MUSICAL

SANTIAGO ALBERTO REYES GONZÁLEZ

Trabajo de Grado para optar al título de Especialista en Estadística Aplicada

Asesor

John Edward Forigua Parra
Coordinador de Posgrados en Ciencias Básicas

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
FACULTAD DE CIENCIAS BÁSICAS
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA
BOGOTÁ, D.C.
2017

Nota de Aceptación

Firma de Presidente del Jurado

Firma del Jurado

Firma del Jurado

Bogotá, 22 de Julio de 2017

Las Directivas de la Fundación Universitaria Los Libertadores, los Jurados Calificadores y el cuerpo Docente no son responsables por los criterios e ideas expuestas en el presente documento. Estos corresponden únicamente al autor.

A Dios por permitirme tener la salud y la posibilidad de cumplir con este objetivo en mi vida.

A mis padres y mi esposa Lucy por su amor y apoyo incondicional y a mi hermano Nicolás por su gran colaboración al permitirme trabajar con sus datos y por apoyarme en lo que necesité.

A los docentes de la Universidad que dedicaron su tiempo a atender mis inquietudes.

Contenido

Introducción	11
Pregunta de Investigación	12
Objetivo General	12
Objetivos Específicos.....	12
Aspectos generales de la Metodología.....	12
Marco Teórico.....	13
Análisis Univariado.....	21
Análisis univariado variable satisfacción.	21
Análisis univariado variable Número de Parpadeos.	25
Análisis univariado variable tiempo de respuesta.	28
Análisis Univariado Variable Latencia del Parpadeo.....	32
Análisis Bivariado	37
Análisis bivariado variable Satisfacción.....	37
Análisis bivariado variable Número de Parpadeos.....	40
Análisis bivariado variable Tiempo de Respuesta.....	43
Análisis bivariado variable Latencia.	45
Conclusiones Análisis bivariado.	46
Conclusiones Generales Análisis Exploratorio	48
Perspectiva No Paramétrica Del Análisis	49
Aplicación de Pruebas de Rangos De Wilcoxon.....	51
Modelos de Elección y Preferencias	55
Modelos de Elección Discreta.....	57
Modelos de elección binaria.	57
Regresión logística.	58
Tipos de Modelos de elección discreta.....	58
Aplicación Del Modelo Logit	59
Razón de aplicación y estructura de datos.....	59
Estadísticos importantes en el Modelo Logit	60
Aplicación de las regresiones.	61
Resultados de las regresiones.	63
Conclusiones de las regresiones.	73
Árboles De Decisión	74
Métodos de crecimiento de los árboles de decisión.	76

Resultados árboles CHAID.	82
Conclusiones de los árboles CHAID.	95
Conclusiones Generales Del Análisis.	97
Referencias.....	103

LISTA DE TABLAS

Tabla 1. Distribución frecuencias variable satisfacción original	22
Tabla 2. Distribución Frecuencias Variable Satisfacción Dicotómica	24
Tabla 3. Estadísticos Descriptivos de la variable Número de Parpadeos	25
Tabla 4 Distribución frecuencias variable número de parpadeos dicotómica	27
Tabla 5 Estadísticos descriptivos variable tiempo de respuesta.	29
Tabla 6 Distribución frecuencias variable tiempo de respuesta dicotómica.	31
Tabla 7 Estadísticos descriptivos de la variable latencia del parpadeo.	32
Tabla 8 Distribución frecuencias variable latencia del parpadeo	34
Tabla 9 Tablas de contingencia satisfacción VS modo y tempo	38
Tabla 10 Pruebas chi cuadrado satisfacción vs modo y tempo.	39
Tabla 11 Medidas de dirección e intensidad. Satisfacción vs modo y tempo.	40
Tabla 12 Tablas de contingencia número de parpadeos vs modo y tempo.	41
Tabla 13 Pruebas chi cuadrado número de parpadeos vs modo y tempo.	42
Tabla 14 Medidas de dirección e intensidad. número de parpadeos vs modo y tempo.	42
Tabla 15 Tablas de contingencia tiempo de respuesta vs modo y tempo.	43
Tabla 16 Pruebas chi cuadrado tiempo de respuesta vs modo y tempo.	44
Tabla 17 Medidas de dirección e intensidad. Número de parpadeos vs modo y tempo.	44
Tabla 18 Tablas de contingencia latencia vs modo y tempo.	45
Tabla 19 Pruebas chi cuadrado latencia vs modo y tempo.	45
Tabla 20 Medidas de dirección e intensidad latencia vs modo y tempo.	46
Tabla 21 Combinaciones para aplicación de pruebas de Wilcoxon	53
Tabla 22 Aplicación prueba signos de rangos de Wilcoxon, valores p	54
Tabla 23 Estadísticas generales promedio regresiones logísticas.	64
Tabla 24 R cuadrado modo y tiempo vs satisfacción.	65
Tabla 25 matriz de clasificación modo y tiempo vs satisfacción.	66
Tabla 26 Significancia variables modo y tiempo vs satisfacción	66
Tabla 27 R cuadrado modo y tiempo vs número de parpadeos	67
Tabla 28 R matriz de clasificación modo y tiempo vs número de parpadeos.	68
Tabla 29 Significancia variables modo y tiempo vs número de parpadeos	69
Tabla 30 R cuadrado Modo y Tempo VS Tiempo de Respuesta.	69
Tabla 31 R matriz de clasificación modo y tiempo VS tiempo de respuesta.	70
Tabla 32 Significancia variables modo y tiempo VS tiempo de respuesta.	71
Tabla 33 R cuadrado modo y tiempo vs latencia	71
Tabla 34 Matriz de clasificación modo y tiempo vs latencia.	72
Tabla 35 Significancia variables modo y tiempo vs latencia.	73
Tabla 36 Comparación de métodos para árboles de decisión	78
Tabla 37 Árboles Chaid modo y tiempo vs satisfacción.	83
Tabla 38 Matriz de clasificación modo y tiempo vs satisfacción	84
Tabla 39 Riesgo modo y tiempo vs satisfacción.	85
Tabla 40 Ganancias modo y tiempo vs satisfacción	86
Tabla 41 Árboles CHAID modo y tiempo vs número de parpadeos	87
Tabla 42 Matriz de clasificación modo y tiempo vs número de parpadeos	88
Tabla 43 Riesgo modo y tiempo vs número de parpadeos	88
Tabla 44 Ganancias modo y tiempo VS número de parpadeos	89
Tabla 45 Árboles CHAID modo y tiempo vs tiempo de respuesta.	90

Tabla 46 Matriz de clasificación modo y tempo vs tiempo de respuesta	91
Tabla 47 Riesgo modo y tempo vs tiempo de respuesta.....	91
Tabla 48 Ganancias modo y tempo vs tiempo de respuesta.	92
Tabla 49 Árboles CHAID modo y tempo vs latencia.....	93
Tabla 50 Matriz de clasificación modo y tempo vs latencia.....	94
Tabla 51 Riesgo modo y tempo vs latencia	94
Tabla 52 Ganancias modo y tempo vs latencia.....	95

LISTA DE FIGURAS

Figura 1. Distribución frecuencias variable satisfacción original.....	22
Figura 2. Distribución frecuencias variable satisfacción dicotómica	24
Figura 3. Distribución de la variable Número de Parpadeos	26
Figura 4. Distribución de la variable número de parpadeos dicotómica.....	27
Figura 5. Distribución de la variable tiempo de respuesta.....	30
Figura 6 Distribución de la variable tiempo de respuesta dicotómica	31
Figura 7. Distribución de la variable latencia del parpadeo.....	33
Figura 8. Distribución de la variable Latencia del Parpadeo.	35

Introducción

La relación entre la música y las emociones ha sido objeto de particular interés para la humanidad desde tiempos remotos, sin embargo, de acuerdo con autores como Benedetti (2002), en muchos de los trabajos en el área de la Psicología, relacionados con la música como manuales y algunos libros, el enfoque se da más hacia los aspectos cognitivos que hacia los emocionales.

De otro lado, no es fácil encontrar estudios en los que se observen análisis estadísticos exhaustivos y no tan descriptivos respecto al tema y es así como se presenta una oportunidad para poder profundizar más en este aspecto y por medio del presente trabajo realizar un análisis comparativo que permita generar alguna conclusión estadística de los efectos emocionales que tienen ciertas variaciones musicales sobre los participantes de la muestra en el experimento piloto realizado.

Los datos provienen de una muestra de 50 personas cada una de las cuales escucha 5 piezas musicales en variaciones de modo mayor y menor y de tempo rápido y lento, de este modo cada participante califica en una escala de satisfacción de 0 a 9 cada una de estas 4 combinaciones en las que también se les registran otras medidas fisiológicas relacionadas con la satisfacción, como el número de parpadeos, el tiempo de respuesta y la duración de los parpadeos o latencia.

Se realiza un análisis estadístico comparativo entre 3 técnicas que por su naturaleza se ajustan al tipo de datos estudiados y además tienen la ventaja de provenir del mundo paramétrico y no paramétrico: Pruebas de Rangos de Wilcoxon, Regresión logística y Árboles de decisión, se generan conclusiones respecto a los efectos que generan las piezas y sus variaciones sobre los participantes y sobre la técnica que mejor se ajuste y explique los resultados.

Pregunta de Investigación

Es así que el presente trabajo se genera la siguiente pregunta de investigación: ¿Existe evidencia estadística de efectos sobre la satisfacción generados por las variables Tempo y Modo Musical?

Objetivo General

Identificar la existencia de evidencia estadística que sustente efectos causales entre las variables Tempo y Modo Musical y las variables Respuesta.

Objetivos Específicos

Evidenciar estadísticamente si hay efectos de causalidad entre las variables Tempo y Modo Musical y la variable respuesta principal Satisfacción.

Evidenciar estadísticamente si hay efectos de causalidad entre las variables Tempo y Modo Musical y las otras variables respuesta Número de Parpadeos, Tiempo de Respuesta y Latencia del Parpadeo.

Aspectos generales de la Metodología

El presente análisis se enmarca dentro del tipo de Investigación Cuantitativo Explicativo (Sampieri, Fernandez & Baptista 2006), teniendo en cuenta la intención de explicar por medio de las herramientas estadísticas más apropiadas los efectos que puedan llegar a tener las variables Modo y Tempo Musical sobre las variables Respuesta Satisfacción, Número de Parpadeos, Latencia del Parpadeo y Tiempo de Respuesta.

De esta forma la metodología utilizada aborda los pasos de Análisis Exploratorio de los Datos de forma univariada y bivariada, con el propósito inicial de entender el comportamiento de los datos tanto en sus variables dependientes como en las independientes, lo que resulta útil

para tener una primera aproximación al comportamiento de las mismas y de si inicialmente se observa algún efecto causal entre las explicativas y las respuesta.

Una vez realizado el análisis exploratorio y descriptivo, se aborda la búsqueda de evidencia estadística de dichos efectos causales tanto desde el punto de vista paramétrico como del no paramétrico utilizando así las técnicas que más se adapten al tipo y relación de los datos analizados. Así, finalmente se generarán las respectivas conclusiones a través de la comparación de los resultados generados en la aplicación de dichas técnicas, con miras a robustecer la existencia o no de evidencia estadística de los mencionados efectos.

Marco Teórico

Teniendo en cuenta que el principal objetivo del presente trabajo será el de entender si hay algún tipo de relación causal entre las variables que se plantean como independientes y sus efectos en las que se plantean como explicadas, será necesario establecer los ejes conceptuales sobre los cuales se abordará dicha problemática. En este sentido en primer lugar se toma como referencia el modelo psicológico propuesto por Scherer y Zetner (1977) para el análisis de los efectos que produce la música sobre las emociones en los seres humanos, modelo que consta de 5 elementos principales agrupados en 2 grandes categorías las estructurales y las no estructurales.

En las estructurales se encuentran en un primer lugar las características que dichos autores llaman “de Segmento” y son elementos puramente acústicos que están sujetos con mayor facilidad a análisis cuantitativos (acordes e intervalos musicales). En segundo lugar, están las que llaman características de Supersegmento con mayor complejidad que las anteriores teniendo en cuenta que corresponden a articulaciones en el tiempo de las mismas (están en este grupo elementos como la Melodía, el Tempo, el Ritmo y la Armonía). Finalmente, en las características no estructurales se encuentran todas aquellas que no tienen

que ver directamente con la música sino más con el contexto, con el oyente mismo y con el músico que interpreta la pieza.

Desde esta perspectiva, el experimento realizado por estos autores es el punto de partida para el analizado en este trabajo que contempla las variables independientes Tempo y Modo y sus efectos en la satisfacción medida directamente y con otras variables dependientes como el Número de parpadeos, la latencia de los mismos y el tiempo de respuesta de los sujetos.

En un segundo lugar el abordaje a la problemática que se realiza en el presente trabajo busca hacerse desde una perspectiva molecular más que molar, esto tomando la idea de una mayor profundización y utilidad práctica futura del entendimiento de las relaciones causales que pueden llegar a tener los elementos musicales y sus variaciones sobre las emociones, perspectiva que Fabrizio Benedetti (2002) conceptualiza y explica en detalle en Efectos de Modo y Tempo sobre el contenido musical, allí el autor menciona la falta de una perspectiva molecular en el abordaje de la problemática de los efectos emocionales generados por la música. En el mencionado artículo el autor parte de la evidente existencia de la relación entre música y emociones y el interés que desde la antigüedad ha generado dicha relación, sin embargo, menciona de forma muy interesante, cómo la psicología moderna por un lado parece haberla ignorado y, por otro, cómo en manuales de psicología de música o de las emociones no se incluyen capítulos de música o se concentran más en aspectos de cognición y percepción puros que en aspectos emocionales propiamente dichos.

Es así como se adopta esta posición conceptual para explicar de la mejor forma los efectos del Modo y Tempo musical y sus variaciones en las variables dependientes utilizadas, como la satisfacción y las otras representantes de la misma desde el punto de vista fisiológico. De esta forma, la visión molecular, que como bien lo menciona Benedetti, implica buscar regularidades en la expresión musical de las emociones a través de la investigación de los

parámetros regularmente manejados por compositores y músicos y “a partir de estas unidades mínimas de música, apreciar los efectos en las personas”

En un primer abordaje a la mencionada problemática desde la óptica no paramétrica se toma como referencia la aplicación de pruebas de Rangos de Wilcoxon que realiza Fabrizio Benedetti en su trabajo Efectos de Modo y Tempo sobre el Contenido Musical. Allí el autor del mencionado trabajo realiza un experimento con 49 sujetos en el que se propone medir los efectos que tiene “la manipulación de dos parámetros musicales de gran importancia -tempo y modo- sobre el contenido musical percibido en una melodía”. Para lo cual construye una página en internet en HTML de la forma más sencilla posible, minimizando al máximo cualquier interferencia en aspectos visuales que pudieran interferir en las emociones de los sujetos, descartando así tecnologías como Flash, hojas de estilo en cascada y colores y fuentes especiales en general. En la página los sujetos son asignados aleatoriamente a 4 piezas musicales creadas en formato MIDI con las variaciones en tempo y modo respectivas, piezas a las que al finalizar su reproducción una única vez, el sujeto asignaba puntuaciones en una escala Likert de 5 puntos según las 4 emociones principales (alegría, tristeza, ira, miedo), datos que se recogían mediante un script que verificaba la completitud de la información.

Según Benedetti (2002), la falta de normalidad en la distribución lo lleva a aplicar la mencionada técnica no paramétrica en la que los resultados, de acuerdo al autor, confirman en una primera instancia la independencia de los parámetros al no mostrar significancia en sus interacciones, lo que facilita su manipulación, mientras que en una segunda instancia las pruebas mostraron la identificación notable de las emociones alegría y tristeza, en donde Tempo mostraba un mayor peso relativo Frente a Modo, en contraste con identificación más pobre de las emociones Miedo e Ira.

Respecto al análisis exploratorio inicial de los datos del análisis se toma la metodología y estadísticos sugeridos por Jaime Fierro en su artículo Análisis univariado, bivariado y

variables de Control, en donde el análisis se realiza de una forma muy completa, como bien su título lo menciona, desde estas dos perspectivas (univariada y bivariada) con el objetivo de entender comportamientos individuales y conjuntos de las variables de interés. En este se tiene en cuenta el tipo de variables para la aplicación de las medidas estadísticas apropiadas según la perspectiva desde la que se está realizando el análisis. Así,

en el análisis univariado en las variables métricas se usan medidas de tendencia central (media, mediana y moda), de dispersión (Varianza y desviación estándar) y localización, etc, mientras que en el de variables nominales distribuciones de frecuencia principalmente. De otro lado, el análisis bivariado trata de identificar un primer acercamiento de los efectos que se producen entre las variables de interés, en este caso nominales, por lo que se utilizan principalmente pruebas de significancia, estadísticos como el coeficiente Chi cuadrado de Pearson, Coeficiente V de Cramer, coeficiente de Spearman, entre otras.

En la elección de las pruebas de rangos de Wilcoxon se tendrá en cuenta conceptos y criterios expuestos por Martha Vázquez Villazón y Maria Eugenia de Bernard (1992) en *Métodos Estadísticos Aplicados a la Investigación Psicológica*, en donde uno de los principales criterios es sobre el conocimiento de los parámetros poblacionales que se utilizan en las técnicas paramétricas, de esta forma uno de los motivos más relevantes para elegir una técnica no paramétrica es el hecho de no conocer los parámetros poblacionales ni su distribución, en segundo lugar un criterio de tamaño muestral en donde se dice que para minimizar los errores de las estimaciones se requiere un número significativo de muestra, lo que en este caso no se da teniendo en cuenta la limitación de recursos. En tercer lugar, el hecho de que los datos analizados no son de intervalo sino nominales ajustándose mejor a métodos no paramétricos. Finalmente, que el análisis se centra en contraste de dos muestras más que en averiguar si la extracción de una muestra proviene de alguna población determinada y además proviene de muestras independientes.

Por otro lado, y con miras a enriquecer el análisis en el presente trabajo, se ha tomado la decisión de utilizar dos técnicas que se ajustan muy bien a la estructura de datos, en la que se busca identificar preferencias del grupo de individuos analizado respecto a sus niveles de satisfacción. Es así que se hace necesario definir los conceptos que han dirigido el análisis al mundo de los modelos de elección discreta, desde el cual se ha tomado la decisión de utilizar la regresión logística como una de las técnicas que se adapta correctamente al objetivo del presente trabajo, esto, no sin antes mencionar la importancia y la decisión de usar también la técnica de árboles de decisión y clasificación teniendo en cuenta su buena adaptabilidad a la estructura del análisis y su ventaja en combinar la flexibilidad del mundo paramétrico y no paramétrico en una sola técnica.

Respecto a los modelos de elección discreta se ha tomado como base la conceptualización que hace Mario Mazzochi (2008) en su trabajo *Statistics for Marketing and Consumer Research*. En este trabajo Mazzochi hace referencia a los dos grupos de teorías que se usan regularmente en los análisis de elección que se realizan para el consumidor. En primer lugar, se encuentran las teorías por preferencias reveladas del consumidor, en donde se busca identificar dichas preferencias de forma indirecta, es decir sin preguntar directamente al consumidor, sino en lugar de ello recolectando información que permita identificarlas, de acuerdo con el autor, desde el punto de vista del esfuerzo considerable que se necesita para recolectar esta información esta metodología resulta más costosa. De otro lado están las teorías de preferencias declaradas que abordan el estudio del consumidor de forma directa, esto es, preguntándole al consumidor sobre sus preferencias. Bajo este último grupo de teorías se encuentran los modelos de elección discreta, ellos provienen de una generalización de la regresión lineal al caso en donde la variable dependiente es no métrica y puede tener valores binarios, ordinales o categóricos y las variables independientes por su lado pueden ser métricas, categóricas o binarias, las variables dependientes no métricas violan los supuestos de

homocedasticidad y normalidad, situación que se soluciona es los modelos de elección discreta a través del uso de funciones de enlace y variables latentes para poder relacionar los comportamientos de la variable explicativa con la dependiente, aspecto que se explicará más en detalle en el correspondiente apartado. Es así que Mazzochi sugiere la regresión logística como uno de los modelos de elección discreta más apropiados para el análisis de preferencias declaradas con variables categóricas, que a través de una función de enlace (función logística) logra crear la técnica para relacionar variables categóricas dependientes con otras categorías independientes, específicamente la regresión logística binaria para variables con dos variaciones o categorías como en el caso que nos atañe en el presente trabajo.

Finalmente, en cuanto a los árboles de decisión, se toma como referencia la definición de los autores Vanesa Berlanga, Maria José Rubio y Ruth Vilà (2013) en su artículo “Cómo Aplicar Árboles de Decisión en SPSS”. De Acuerdo con ellos, esta técnica se puede usar como una herramienta muy flexible que sirve para la segmentación, la estratificación, la predicción, la reducción de datos y el filtrado de variables, la identificación de interacciones, la fusión de categorías y la discretización de variables continuas, de acuerdo con ellos estos árboles crean un modelo de clasificación basado en un diagrama de flujo en donde se clasifican casos en grupos o pronostican valores de una variable dependiente con base en los valores de las independientes o variables predictoras. Como se puede ver es una técnica que se adecúa muy bien al análisis que se realiza en este trabajo al permitir la predicción y el trabajo con variables categóricas como lo son el tempo, modo musical y las variables explicadas usadas (Número de Parpadeos, Satisfacción, Latencia y tiempo de respuesta dicotómicas).

En el mencionado artículo se explican los 4 métodos de división que utilizan los árboles y que corresponden a las siguientes:

- CHAID (Chi-square automatic interaction detector): lo definen como algoritmo de árbol estadístico que explora datos de forma rápida y eficaz y que crea segmentos y

perfiles respecto al resultado deseado. El proceso que realiza el algoritmo permite la detección automática de interacciones mediante pruebas Chi cuadrado en donde en cada paso elige la variable independiente predictora que presenta la interacción más fuerte según la variable dependiente que se esté trabajando, también en cada paso, la explicación en detalle del mismo se realiza en el correspondiente apartado en donde se selecciona esta técnica como la más apropiada (ver página 82).

- CHAID exhaustivo: Es una modificación del algoritmo CHAID que trata a todas las variables por igual independientemente del tipo y número de categorías.
- Árboles de Clasificación y regresión (CRT-Classification and regression trees): Técnica que consiste en un algoritmo de árbol binario que hace particiones de datos y genera subconjuntos precisos y homogéneos. El Algoritmo CRT divide los datos en segmentos para que sean lo más homogéneos posible respecto a la variable dependiente.
- QUEST (Quick Unbiased, efficient statistical tree): Consiste en un algoritmo estadístico que selecciona variables sin sesgo y crea árboles binarios precisos de forma rápida.

Como se puede ver de por sí los árboles de decisión son técnicas flexibles y completas que permiten realizar análisis de datos categóricos y continuos al incluir procedimientos estadísticos inferenciales que garantizan la significancia de los grupos creados en las ramas de los árboles y que finalmente generan la predicción. Una vez entendida la bondad de estas técnicas, solo queda seleccionar el método de crecimiento que más se ajuste a los datos del presente análisis, encontrando así la definición de la técnica CHAID que realiza Carlo Magno Araya Alpizar (1994) en su artículo “Segmentación de Mercados usando la técnica CHAID”, definición que facilita la selección de este algoritmo como el método de crecimiento más adecuado de acuerdo a los objetivos del presente análisis.

Según Carlo Magno la técnica CHAID tiene su fundamento en las técnicas “Automatic Interaction Detection” AID y Classification and Regression Trees (CART). El autor define la técnica como una especie de regresión múltiple para variables nominales, ordinales, categóricas, discretas y discontinuas, en la que existe una variable dependiente y al menos una variable independiente, que trata de predecir la variable respuesta a través de las variables predictoras. También explica como un análisis de regresión tradicional a diferencia de la técnica CHAID, trabaja con variables continuas que pueden ser de intervalo o de razón. Además de lo anterior, como bien lo dice Carlo Magno, el CHAID ofrece las bondades de trabajar con múltiples variables categóricas en un algoritmo automático que utiliza pruebas Chi cuadrado para identificar las variables y categorías pertinentes que explican la variable independiente, ahorrando así bastante tiempo al investigador, al evitar que éste realice múltiples tabulaciones cruzadas sin encontrar alguna relación importante entre las variables, es así que se decide seleccionar dicha técnica para la realización del respectivo análisis, en paralelo con la regresión logística explicada anteriormente, análisis en los que se profundizará en los respectivos apartados.

Análisis Exploratorio

Con el objetivo de lograr una primera aproximación a los datos y de entender su distribución en las diferentes variables tratadas, tanto de forma univariada como bivariada, se procedió con el respectivo análisis exploratorio del cual salieron las correspondientes gráficas y medidas estadísticas necesarias para alcanzar dicho entendimiento inicial. Así, en primer lugar, se procede con un análisis descriptivo univariado de las variables dependientes en su forma original y dicotomizadas, posteriormente se procede con el análisis bivariado por medio de los cruces de las variables dependientes e independientes que ayudan a conocer si inicialmente existe la relación causal esperada y su magnitud y dirección en caso de ser así.

Para las variables dependientes que en su forma natural son métricas, como el número de parpadeos, el tiempo de respuesta y la duración de los parpadeos, se utilizaron medidas de tendencia central, dispersión y localización para el caso del análisis univariado. Por otro lado, para el análisis de las variables dicotomizadas y la de Satisfacción en su forma original que son variables no métricas, se utilizaron las técnicas descriptivas apropiadas a su naturaleza tal como los son el análisis de frecuencia y los Figuras de barras para el caso univariado y coeficientes Ji Cuadrado de Pearson, Coeficiente V de Cramer y de Correlación de Spearman para el caso del bivariado.

Análisis Univariado

Como se mencionó anteriormente, el análisis univariado que se realiza busca entender el comportamiento de cada una de las variables dependientes analizadas tanto en su forma original como en la dicotómica, esto permitirá entender la forma en cómo se transformaron las variables originales para construir las variables dependientes que permiten usar las técnicas de análisis posteriores y a la vez entender el comportamiento de distribución de las mismas.

Análisis univariado variable satisfacción.

- Variable de Satisfacción Original

La variable dependiente Satisfacción está medida en una escala del 0 al 9 originalmente para que cada uno de los individuos califique el grado en el que le agrada la pieza musical que escucha, como se puede ver cada uno de los valores representa un atributo cualitativo y son valores discretos por lo que la variable de Satisfacción original es nominal y debe analizarse con distribuciones de frecuencia al igual que la variable transformada de satisfacción dicotómica.

Tabla 1. Distribución frecuencias variable satisfacción original

valores	Pieza 1			Pieza 2			Pieza 3			Pieza 4			Pieza 5		
	Frecuencia	%	% acumulado	Frecuencia	%	% acumulado	Frecuencia	%	% acumulado	Frecuencia	%	% acumulado	Frecuencia	%	% acumulado
0	18	9,0	9,0	26	13,0	13,0	16	8,0	8,0	28	14,0	14,0	15	7,5	7,5
1	23	11,5	20,5	22	11,0	24,0	22	11,0	19,0	18	9,0	23,0	16	8,0	15,5
2	14	7,0	27,5	20	10,0	34,0	28	14,0	33,0	25	12,5	35,5	22	11,0	26,5
3	25	12,5	40,0	20	10,0	44,0	21	10,5	43,5	35	17,5	53,0	19	9,5	36,0
4	17	8,5	48,5	18	9,0	53,0	14	7,0	50,5	20	10,0	63,0	20	10,0	46,0
5	21	10,5	59,0	19	9,5	62,5	23	11,5	62,0	17	8,5	71,5	26	13,0	59,0
6	16	8,0	67,0	21	10,5	73,0	19	9,5	71,5	19	9,5	81,0	16	8,0	67,0
7	21	10,5	77,5	23	11,5	84,5	20	10,0	81,5	12	6,0	87,0	26	13,0	80,0
8	21	10,5	88,0	14	7,0	91,5	14	7,0	88,5	17	8,5	95,5	23	11,5	91,5
9	24	12,0	100,0	17	8,5	100,0	23	11,5	100,0	9	4,5	100,0	17	8,5	100,0
Total	200	100,0		200	100,0		200	100,0		200	100,0		200	100,0	

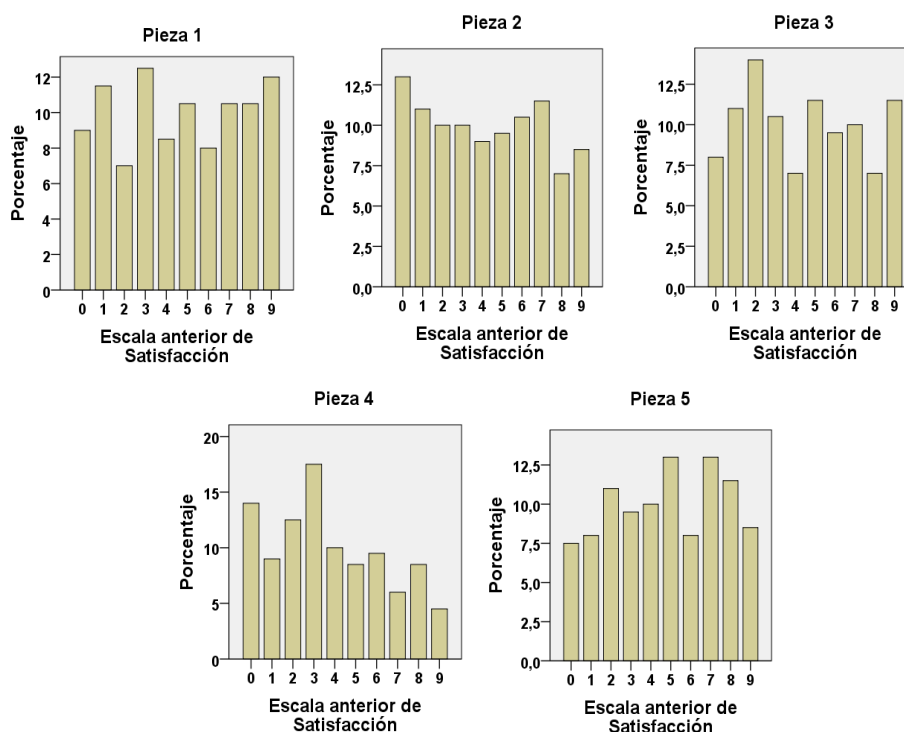


Figura 1. Distribución frecuencias variable satisfacción original

En primera instancia se puede observar que, en general en la mayoría de las piezas musicales, la distribución de frecuencias no muestra concentraciones atípicas en valores o grupos de valores específicos y por el contrario se distribuye de forma casi similar entre los diferentes valores, sin embargo, algunas de las piezas musicales muestran comportamientos un

poco más marcado respecto a las demás según su concentración de frecuencias, como es el caso de las piezas 4 y 5.

Respecto a la pieza 4 se evidencia una concentración marcada en los valores bajos de la escala en comparación a las demás piezas musicales ya que para el valor 4 de la escala ya se encuentra concentrado el 63% de los casos, por el contrario, la pieza 5 muestra el comportamiento totalmente opuesto en donde para el valor 4 el porcentaje acumulado de frecuencia es del 46%, lo que en un primer acercamiento puede mostrar que los niveles de satisfacción con la pieza 4 inicialmente en términos generales son bajos respecto a las demás piezas y los de la 5 los más altos.

Los Figuras confirman lo anteriormente mencionado respecto a tendencias marcadas en las piezas 4 y 5 mostrando visualmente las concentraciones de frecuencias en el lado izquierdo del eje de escala de satisfacción en el primer caso y al lado derecho para el último, sin embargo, para las otras piezas las concentraciones no muestran tendencias tan marcadas o claras como en el caso de las dos piezas mencionadas.

- Variable de Satisfacción Dicotómica

Una vez construida la variable de satisfacción dicotómica que tiene los valores 0 y 1 agrupando los valores más bajos (0 a 4) de la escala original en el primero y los más altos (5 a 9) en el segundo, se procede con el análisis univariado de la misma, esto permite apreciar con más detalle las tendencias de las otras piezas que no se ven tan claras en el análisis de la variable original de Satisfacción.

Tabla 2. Distribución Frecuencias Variable Satisfacción Dicotómica

Pieza	Categorías	Frecuencia	Porcentaje	Porcentaje acumulado
P1	Insatisfecho	97	48,5	48,5
	Satisfecho	103	51,5	100,0
	Total	200	100,0	
P2	Insatisfecho	106	53,0	53,0
	Satisfecho	94	47,0	100,0
	Total	200	100,0	
P3	Insatisfecho	101	50,5	50,5
	Satisfecho	99	49,5	100,0
	Total	200	100,0	
P4	Insatisfecho	126	63,0	63,0
	Satisfecho	74	37,0	100,0
	Total	200	100,0	
P5	Insatisfecho	92	46,0	46,0
	Satisfecho	108	54,0	100,0
	Total	200	100,0	

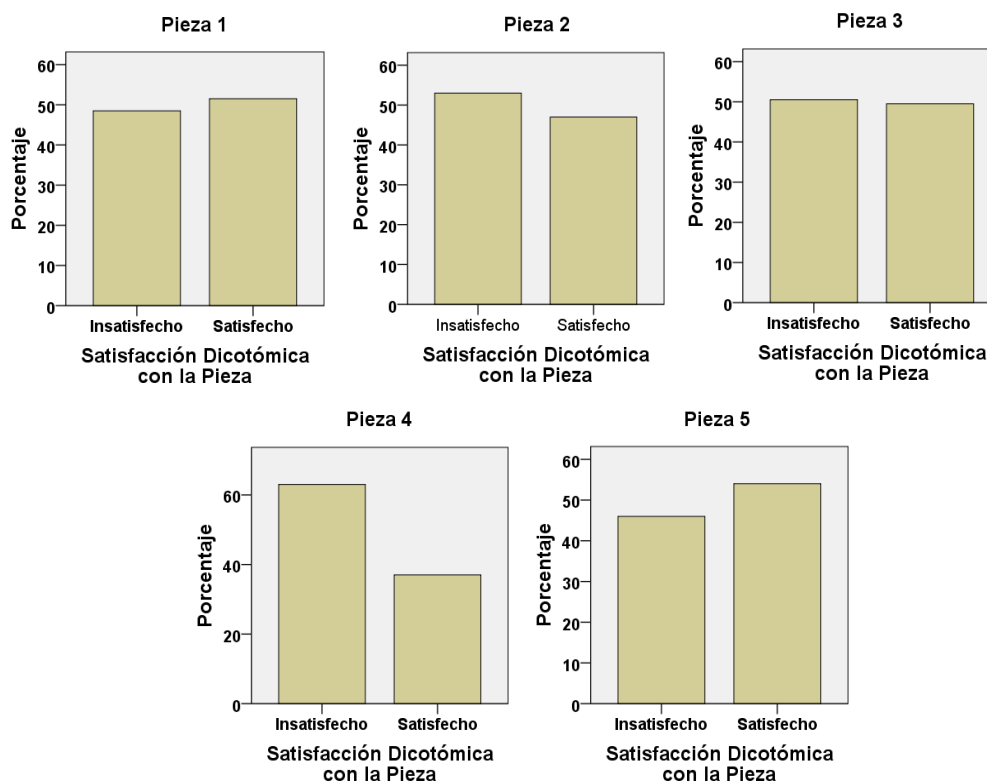


Figura 2. Distribución frecuencias variable satisfacción dicotómica

De acuerdo a lo que se aprecia con la variable dicotómica de Satisfacción, efectivamente las piezas con comportamientos más extremos son las 4 y 5, sin embargo,

también se puede apreciar que la 3 no muestra casi ninguna concentración hacia ningún valor de la variable dependiente dicotómica, mientras que el comportamiento de la pieza 1 se acerca más al de la 5 en donde se registran niveles mayores de satisfacción y el de la pieza 2 se acerca más al de la 4 en donde los niveles de satisfacción registrados son menores.

Análisis univariado variable Número de Parpadeos.

- Variable de Número de Parpadeos en su forma original.

La variable dependiente Número de Parpadeos en su forma original pertenece a un tipo de variable diferente a la de Satisfacción ya que es una variable métrica discreta, de este modo el análisis que le corresponde es diferente al de esta última y aunque permite el análisis de frecuencias por su forma discreta también admite medidas propias de variables métricas como lo son las de tendencia central, dispersión y posicionamiento.

Tabla 3. *Estadísticos Descriptivos de la variable Número de Parpadeos*

Estadísticos	Numero de Pieza					
	P1	P2	P3	P4	P5	
Media	6,35	4,60	4,43	3,71	3,86	
Mediana	4,80	3,53	3,30	2,67	3,00	
Moda	,00	1,76	,00	1,43	,00	
Desv. típ.	5,29	3,92	3,85	3,12	3,33	
Coefficiente Variación	83,21	85,26	86,92	84,19	86,32	
Varianza	27,96	15,38	14,81	9,74	11,08	
Asimetría	,60	,74	,67	,71	,82	
Error típ. de asimetría	,17	,17	,17	,17	,17	
Curtosis	-,90	-,71	-,90	-,81	-,59	
Error típ. de curtosis	,34	,34	,34	,34	,34	
Rango	16,80	12,35	12,00	10,00	10,77	
Mínimo	,00	,00	,00	,00	,00	
Máximo	16,80	12,35	12,00	10,00	10,77	
Percentiles	25	2,22	1,56	1,12	1,33	1,41
	50	4,80	3,53	3,30	2,67	3,00
	75	11,05	7,06	7,80	6,33	5,53

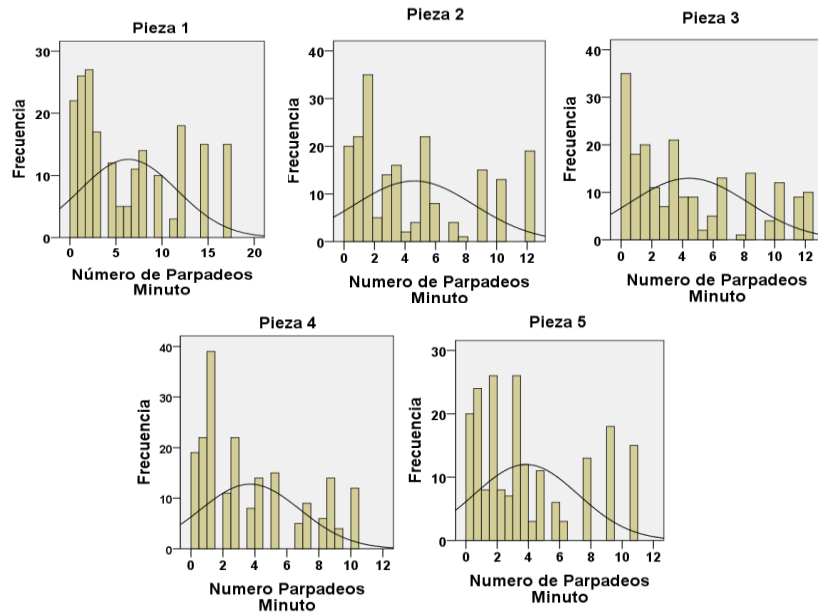


Figura 3. Distribución de la variable Número de Parpadeos

Como se puede apreciar en las gráficas, en general para las 5 piezas musicales la concentración de frecuencias se da hacia los valores más bajos de la variable dependiente en su forma natural, lo que a su vez se puede corroborar al observar las medidas de localización en donde el percentil 75 de la mayoría de las piezas está entre los 5 y 8 parpadeos aproximadamente mientras que los valores máximos se encuentran entre los 10 y 12 parpadeos, lo que sugiere que el 75% de los datos está concentrado en valores relativamente bajos como se menciona anteriormente. Este comportamiento es observado en las piezas musicales 2 a la 5, sin embargo, la pieza 1 se muestra un comportamiento distinto a las demás ya que tiene un rango más amplio y también una acumulación del 75% de los datos en un valor relativamente más alto, sugiriendo una tendencia de concentración más hacia valores altos de la variable dependiente. De otro lado también se aprecian valores del coeficiente de asimetría positivos lejanos de 0 en concordancia con distribuciones de la variable dependiente concentradas en valores bajos como ya se había dicho.

Finalmente, respecto al agrupamiento general de la curva de distribución de cada una de las variables Número de Parpadeos en las 5 piezas, observamos índices de curtosis negativos

lo que indica que los datos tienden a estar más dispersos que agrupados, esto también se puede corroborar con los valores elevados de los coeficientes de variación observados que se encuentran sobre el 83% en todos los casos y con valores de la media que en general se encuentran por encima de todas las medianas.

- Variable de Número de Parpadeos Dicotómica

Tabla 4 *Distribución frecuencias variable número de parpadeos dicotómica*

Pieza	Categorías	Frecuencia	Porcentaje	Porcentaje acumulado
P1	Pocos	92,0	46,0	46,0
	Hartos	108,0	54,0	100,0
	Total	200,0	100	
P2	Pocos	114,0	57,0	57,0
	Hartos	86,0	43,0	100,0
	Total	200,0	100	
P3	Pocos	121,0	60,5	60,5
	Hartos	79,0	39,5	100,0
	Total	200,0	100	
P4	Pocos	125,0	62,5	62,5
	Hartos	75,0	37,5	100,0
	Total	200,0	100	
P5	Pocos	131,0	65,5	65,5
	Hartos	69,0	34,5	100,0
	Total	200,0	100	

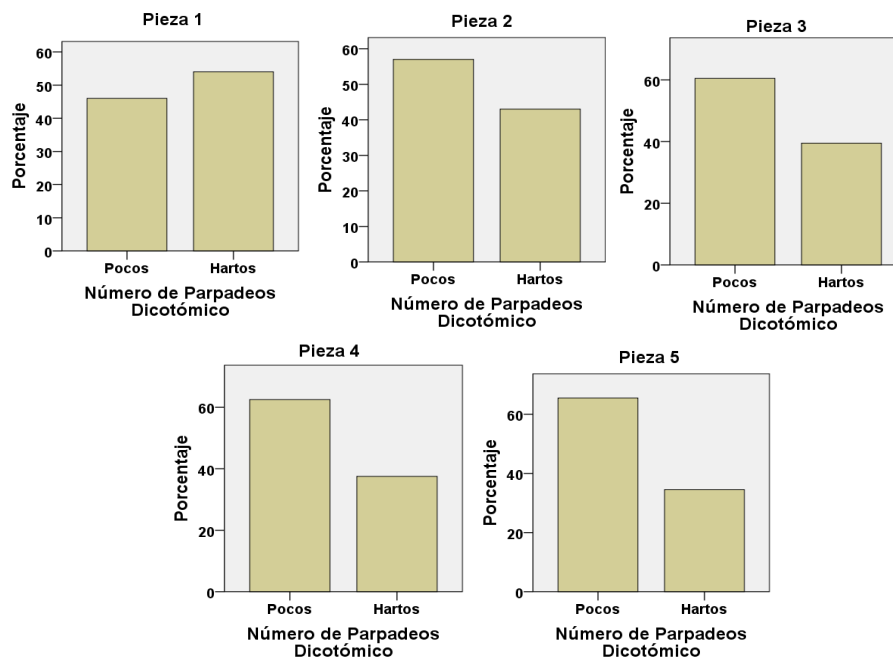


Figura 4. Distribución de la variable número de parpadeos dicotómica

La división de la variable número de parpadeos en hartos y pocos corresponde a medidas fisiológicas en las que las tasas promedio de parpadeo de una persona se encuentran alrededor de 10 parpadeos por minuto, sin embargo, cuando esta se encuentra mirando un objeto fijo como es el caso de un computador estas tasas descienden a niveles de 4 parpadeos por minuto Redmond DE Jr. (1999).

Ya entendiendo la creación de la variable dicotómica de número de parpadeos a través de un criterio general, se observan las distribuciones de la misma y se observa que en la mayoría de los casos la frecuencia se concentra en la categoría “Pocos”, (piezas de la 2 a la 5), lo anterior corrobora las distribuciones de la variable original, como se aprecia en el anterior apartado, sin embargo, es importante ver que en la pieza 1 las distribuciones muestran una forma diferente a las del resto concentrando sus frecuencias en la categoría “Hartos” de la variable dicotómica, de otro lado, la Tabla 4 revela a las piezas 4 y 5 como las más marcadas respecto a este comportamiento de concentración en la categoría “Pocos” con valores de 65% y 65% respectivamente .

Análisis univariado variable tiempo de respuesta.

La variable Tiempo de Respuesta original, corresponde a una variable métrica continua que, a diferencia de la anterior, con escala discreta puede manejar histograma. De acuerdo con teorías del condicionamiento que usa al autor del experimento piloto (2015), esta variable se postula como variable independiente que según el aspecto fisiológico puede representar niveles de satisfacción del individuo con la pieza musical, dichas teorías no son del alcance del presente trabajo, sin embargo se mencionará a modo general que provienen de autores como Michael Domjan (2010) que estudian el condicionamiento clásico y las latencias o tiempos de respuesta como eventos satisfactorios para el individuo estudiado.

- Variable Tiempo de Respuesta en su forma original.

Similar en algunas medidas al caso anterior, esta variable en su forma original, puede ser analizada con medidas de tendencia central, posición y dispersión, sin embargo, como se mencionaba anteriormente difiere de las anteriores en ser una variable continua, que se asemeja más a la siguiente (Latencia del parpadeo).

Tabla 5 *Estadísticos descriptivos variable tiempo de respuesta.*

Estadísticos	Numero de Pieza					
	P1	P2	P3	P4	P5	
Media	9,74	10,60	9,17	9,36	10,02	
Mediana	6,99	8,70	6,64	6,92	7,52	
Moda	8,98	8,98	8,98	8,98	8,98	
Desv. típ.	7,87	7,82	7,73	7,48	7,67	
Coefficiente Variación	80,86	73,79	84,30	79,85	76,60	
Varianza	61,98	61,19	59,72	55,91	58,90	
Asimetría	1,47	1,29	1,65	1,55	1,33	
Error típ. de asimetría	,17	,17	,17	,17	,17	
Curtosis	1,20	,97	1,93	1,80	,70	
Error típ. de curtosis	,34	,34	,34	,34	,34	
Rango	31,53	32,00	31,60	31,52	31,40	
Mínimo	1,62	1,63	1,63	1,61	1,67	
Máximo	33,16	33,63	33,23	33,13	33,07	
Percentiles	25	4,91	5,25	4,16	3,99	4,62
	50	6,99	8,70	6,64	6,92	7,52
	75	11,02	13,92	10,13	11,86	11,57

Como puede verse en la Tabla 5 la dispersión es cercana al 80% en la mayoría de las piezas musicales, en comparación con la variable anterior esta dispersión es menor, también se observa que en todos los casos hay una tendencia de concentración en los valores bajos de la variable coherente con percentiles 75 bastante alejados de los valores máximos y con valores de asimetría positivos. Comparativamente la concentración en esta variable es menor a la de la anterior variable si se observa que los valores de curtosis en este caso son todos positivos.

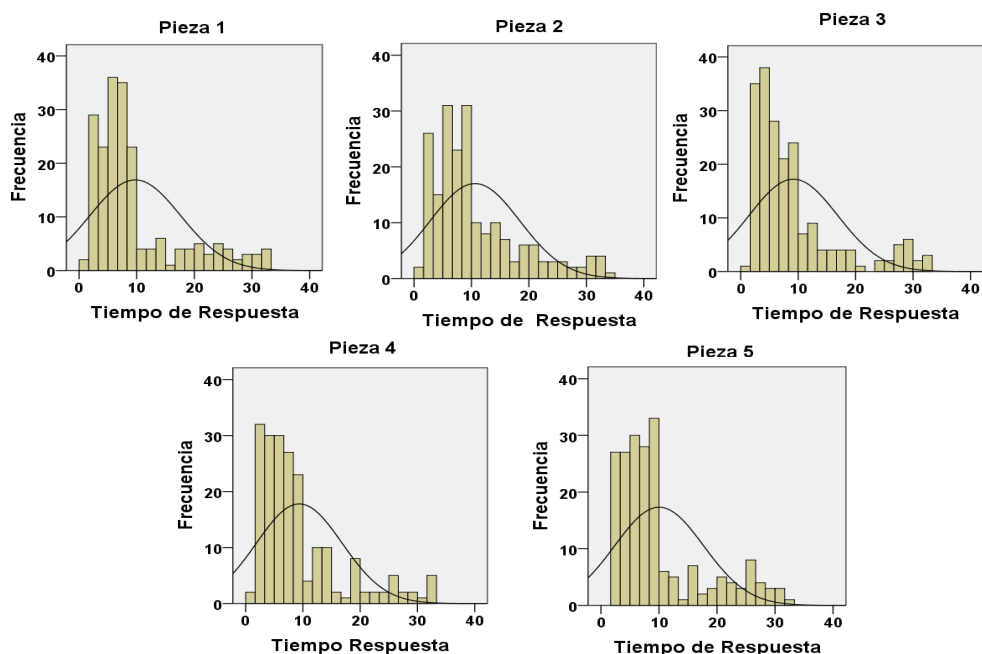


Figura 5. Distribución de la variable tiempo de respuesta

Corroborando los datos contenidos en la Tabla 5, los Figuras exponen las concentraciones hacia la derecha del eje x de las curvas de distribución en las diferentes piezas musicales, como también una menor dispersión respecto a la variable anterior, también se observa una tendencia similar entre todas las piezas en contraste con la anterior variable en donde todas las piezas no comparten la misma tendencia.

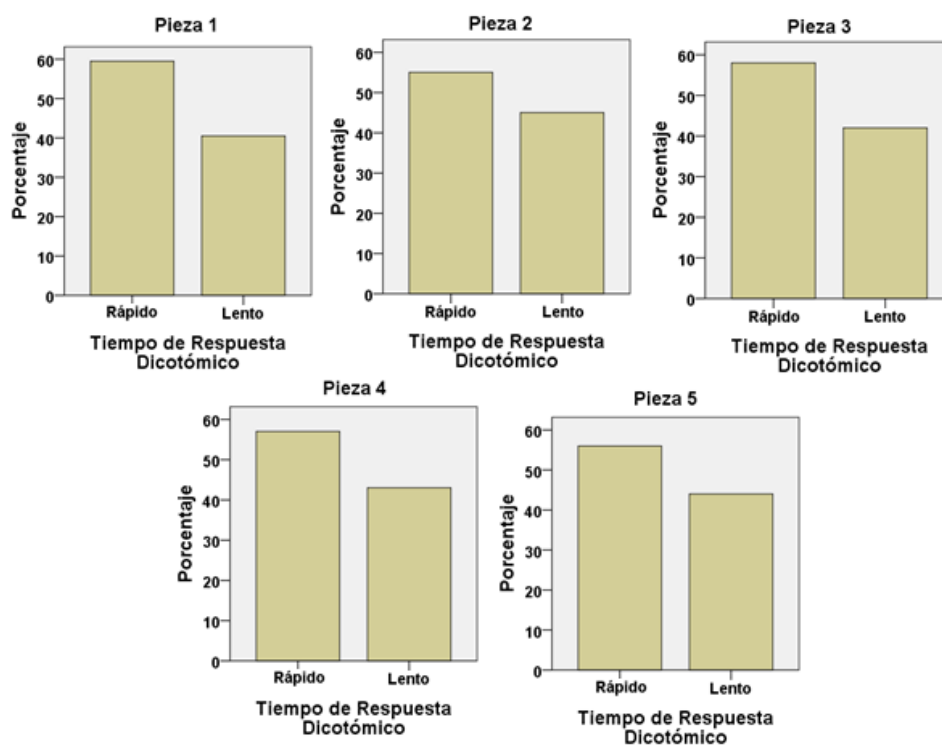
- Variable de Tiempo de Respuesta dicotómica.

La dicotomización de esta variable se hace respecto a la media de cada una de las piezas musicales en términos de los segundos que demoran los individuos en registrar los datos de sus diferentes respuestas

A continuación, al igual que en los casos anteriores se procederá a mostrar la tabla y Figuras de distribución de frecuencia de la variable dicotómica Tiempo de Respuesta Dicotómica.

Tabla 6 *Distribución frecuencias variable tiempo de respuesta dicotómica.*

Pieza	Categorías	Frecuencia	Porcentaje	Porcentaje acumulado
P1	Rápido	119,0	59,5	59,5
	Lento	81,0	40,5	
	Total	200,0	100	
P2	Rápido	110,0	55,0	55
	Lento	90,0	45,0	
	Total	200,0	100	
P3	Rápido	116,0	58,0	58
	Lento	84,0	42,0	
	Total	200,0	100	
P4	Rápido	114,0	57,0	57
	Lento	86,0	43,0	
	Total	200,0	100	
P5	Rápido	112,0	56,0	56
	Lento	88,0	44,0	
	Total	200,0	100	

Figura 6 *Distribución de la variable tiempo de respuesta dicotómica*

Como muestra el Figura 6 y como se mencionaba anteriormente, todas las piezas musicales concentran sus frecuencias en la categoría “Rápido” de la variable Dicotómica en especial la pieza 1 que sobresale de las demás por su marcada concentración en dicha categoría.

Análisis Univariado Variable Latencia del Parpadeo.

Al igual que en los casos anteriores, como ya se mencionaba, la variable Latencia o Duración del parpadeo se configura dentro de las variables respuesta a contrastar ya que, según clasificaciones en el ámbito psicológico específicamente en el campo conductual, representan efectos en la satisfacción de los individuos y sus repuestas a estímulos de acuerdo con la temporalidad de los mismos (Domjan 2010).

- Variable Latencia del Parpadeo en su forma original.

Esta variable se clasifica dentro de las métricas continuas, de esta forma se procederá también con el respectivo análisis de medidas de dispersión, tendencia central y localización con el objetivo de tener un primer acercamiento con esta variable y su comportamiento.

Tabla 7 Estadísticos descriptivos de la variable latencia del parpadeo.

Estadísticos	Numero de Pieza					
	P1	P2	P3	P4	P5	
Media	323	342	346	345	363	
Mediana	200	233	233	217	233	
Moda	200	200	200	167	200	
Desv. típ.	317	308	328	320	325	
Coefficiente Variación	98	90	95	93	89	
Varianza	100.246	95.163	107.383	102.492	105.338	
Asimetría	1,12	,98	1,04	1,07	,92	
Error típ. de asimetría	,17	,17	,17	,17	,17	
Curtosis	,07	-,23	-,20	-,09	-,43	
Error típ. de curtosis	,34	,34	,34	,34	,34	
Rango	1127	1130	1133	1130	1126	
Mínimo	,00	,00	,00	,00	,00	
Máximo	1127	1130	1133	1130	1126	
Percentiles	25	94	111	110	102	117
	50	200	233	233	217	233
	75	474	549	549	528	614

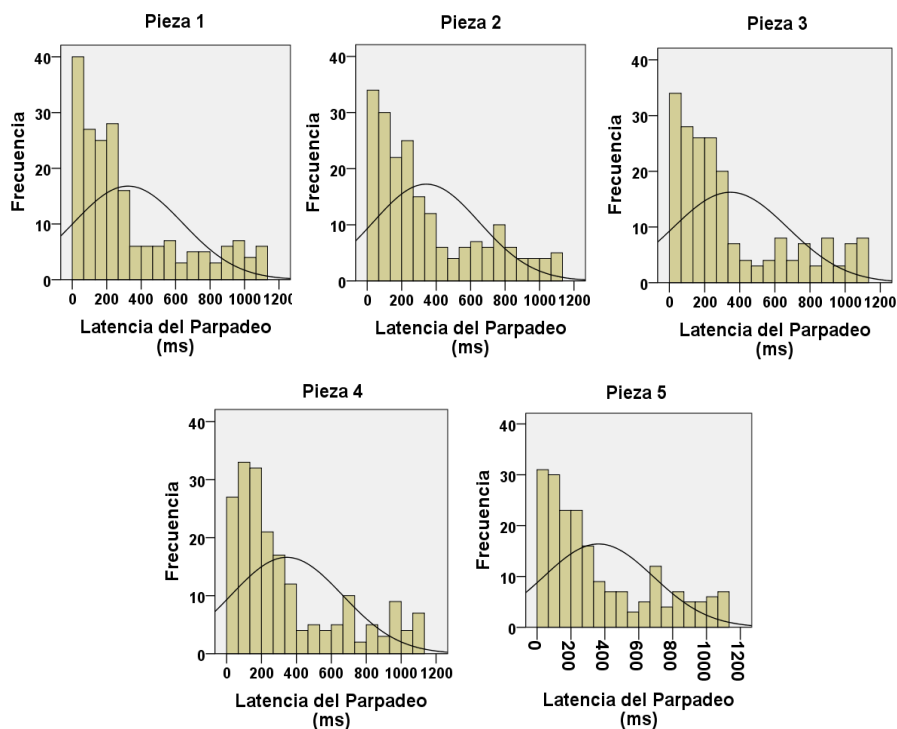


Figura 7. Distribución de la variable latencia del parpadeo

Según la Tabla y Figura 7, las concentraciones de casos en las diferentes piezas musicales se presentan en general hacia los valores bajos de la variable respuesta analizada. De este modo la Tabla 7 muestra valores de asimetría positivos para todos los casos generando curvas de distribución con colas largas en su lado derecho, también presenta valores relativamente bajos de los percentiles 50 y 75 recogiendo así el 50% de los casos en el valor aproximado de 200 y el 75% en el valor aproximado de 500, lo que corrobora dicha concentración hacia la izquierda del eje de la latencia si se tiene en cuenta que los rangos que muestra la tabla se encuentran alrededor de los 1100 ms (milisegundos). Finalmente, la curtosis negativa que se presenta en la mayoría de los casos (Piezas de la 2 a la 5) evidencia niveles de agrupación general menores a los de una curva normal con colas más cortas, sin embargo, el caso de la pieza 1 se sale de este contexto ya que con curtosis

levemente positiva representa una curva de distribución más concentrada que la de una distribución normal. De otro lado

- Variable Latencia del Parpadeo Dicotómica.

La variable original de Latencia del parpadeo está tomada en milisegundos, para realizar la transformación de esta variable a la dicotómica se sacan los promedios de cada una de las piezas musicales y de acuerdo con la comparación de cada uno de los valores con dichos promedios se clasifica el registro como “Poca” en caso de estar por debajo y “Harta” en caso de estar sobre el mismo.

Tabla 8 *Distribución frecuencias variable latencia del parpadeo*

Pieza	Categorías	Frecuencia	Porcentaje	Porcentaje acumulado
P1	Poca	125,0	62,5	62,5
	Harto	75,0	37,5	
	Total	200,0	100	
P2	Poca	120,0	60,0	60
	Harto	80,0	40,0	
	Total	200,0	100	
P3	Poca	124,0	62,0	62
	Harto	76,0	38,0	
	Total	200,0	100	
P4	Poca	120,0	60,0	60
	Harto	80,0	40,0	
	Total	200,0	100	
P5	Poca	117,0	58,5	58,5
	Harto	83,0	41,5	
	Total	200,0	100	

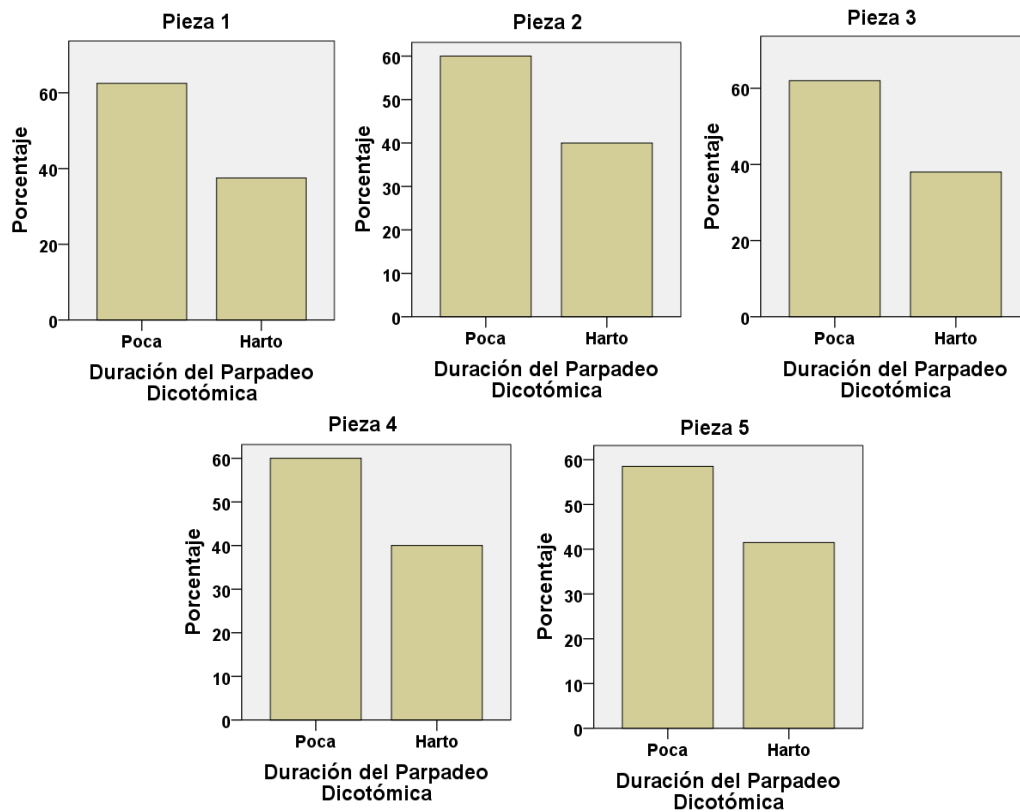


Figura 8. Distribución de la variable Latencia del Parpadeo.

Tal y como se puede observar en la gráfica y tabla 8 la concentración de las frecuencias se hace presente en la categoría “poca” de la variable Latencia del Parpadeo Dicotómica, siendo las piezas 1 y 3 las que mayor concentración presentan en esta categoría, mientras que por el contrario la número 5 muestra el comportamiento de menor concentración en esta categoría respecto a las demás piezas.

Conclusiones análisis univariado

El análisis univariado muestra un primer acercamiento importante para poder ir entendiendo el comportamiento de las variables respuesta analizadas. En primer lugar, se puede apreciar que dichas variables presentan coeficientes de variación mayores al 80%, sin embargo, a pesar de dicha variabilidad, se observa presencia de agrupación de los datos teniendo en cuenta que todas las variables con excepción de Tiempo de Respuesta, cuentan con valores de curtosis negativos. Finalmente, se hace evidente la presencia de asimetría con coeficientes

positivos en todas las variables dependientes, lo que resulta consistente con curvas concentradas, que en este caso agrupan valores en la escala más baja de las variables respuesta analizadas.

Respecto a las piezas musicales, se observan tendencias marcadas y sobresalientes en la 1 y la 5 respecto a las variables dependientes que se analizan. Las variables Tiempo de Respuesta y Latencia del Parpadeo muestran comportamientos similares concentrando la mayoría de sus casos en los valores “Rápido” y “Poca” respectivamente, se aprecia también que en ambas variables la pieza 1 se caracteriza por ser la más representativa de este comportamiento. Por otro lado, en las variables Satisfacción y Número de Parpadeos, todas las piezas musicales no muestran el mismo comportamiento entre sí, en el caso de la primera, las piezas 1 y 5 concentran más del 50% de sus casos en la Categoría “Satisfecho” de acuerdo a la variable dicotómica, mientras que el resto de las piezas lo hacen en la categoría “Insatisfecho”, para el caso de la segunda, el comportamiento de la pieza 1 difiere de las del resto ya que concentra más del 50% de sus casos en la categoría “Hartos”, mientras que el resto de las piezas musicales lo hacen en la categoría “Pocos” de la variable dicotómica, siendo en este último caso la pieza 5 la más representativa de dicho comportamiento.

Finalmente en términos de Satisfacción a primera vista la pieza 1 parece ser la única que mantiene consistentemente resultados similares teniendo en cuenta que en la variable de satisfacción original concentra la mayoría de sus casos en la categoría “Satisfecho” y en las demás variables específicamente en las de Latencia del Parpadeo y Tiempo de Respuesta se comporta con resultados en la misma dirección siendo la que más concentra sus casos en las categorías “Poca” y “Rápido” respectivamente, por otro lado, la pieza 5 a pesar de que también concentra sus casos en la categoría “Satisfecho” de la variable de Satisfacción dicotómica no conserva la misma línea en sus resultados para las otras variables de forma tan marcada como lo hace la 1, ya que si bien si concentra sus casos en las categorías “Rápido” y “Poca” de las

variables Tiempo de Respuesta y Latencia del parpadeo dicotómicas, no es de las más representativas de este comportamiento.

Análisis Bivariado

El análisis bivariado es el segundo y último paso a seguir en el análisis exploratorio y es el que permitirá identificar con más claridad los comportamientos de causalidad que puedan existir entre las variables dependientes y las explicativas ya que consiste en generar los estadísticos que salen del cruce entre estos dos tipos de variables según su rol.

En este análisis se harán las respectivas tablas de contingencia entre las variables explicativas Modo y Tempo Musical y cada una de las variables dependientes dicotómicas analizadas, también se generarán las tablas con las respectivas pruebas Chi cuadrado para entender la asociación entre las variables respuesta y las explicativas, como también las tablas con estadísticos como Coeficiente V de Cramer, de Contingencia que ayudarán a entender la intensidad y dirección de las asociaciones respectivamente.

Análisis bivariado variable Satisfacción.

A continuación se mostrarán los estadísticos y tablas correspondientes al análisis bivariado que permitirá entender si hay relación entre las variables independientes Tempo y Modo Musical, con la variable dependiente Satisfacción Dicotómica, de igual modo en el caso de existir alguna relación, dichas tablas y estadísticos permitirán también observar y entender la dirección e intensidad de la misma.

Tabla 9 Tablas de contingencia satisfacción VS modo y tempo

Tempo															
Categorías	P1			P2			P3			P4			P5		
	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total
Insatisfecho	60%	37%	49%	68%	38%	53%	62%	39%	51%	77%	49%	63%	59%	33%	46%
Satisfecho	40%	63%	52%	32%	62%	47%	38%	61%	50%	23%	51%	37%	41%	67%	54%
Modo															
Categorías	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total
Insatisfecho	70%	27%	49%	62%	44%	53%	66%	35%	51%	71%	55%	63%	62%	30%	46%
Satisfecho	30%	73%	52%	38%	56%	47%	34%	65%	50%	29%	45%	37%	38%	70%	54%

Como se puede observar en la tabla en todas las piezas musicales parece evidente la relación que se espera entre las variables independientes Tempo y Modo Musical ya que la distribución de los casos, en su mayoría, es siempre mayor al 60% en los cruces entre las categorías Lento e Insatisfecho y Rápido y Satisfecho para la variable Tempo Musical, de igual forma se hace evidente el mismo comportamiento para la variable Modo Musical VS la de Satisfacción en donde se aprecia que en la mayoría de los casos la concentración en los cruces de las categorías Menor con Insatisfecho y Mayor con Satisfecho es mayor al 60%, sin embargo, se observa mayor variabilidad de la distribución en este último caso donde los valores no se sostienen alrededor de un valor en particular, sino que tienden a oscilar en un rango más amplio si se tiene en cuenta que por ejemplo en la pieza 4 la concentración es de 45% mientras que en las piezas como la 1 y la 5, ronda alrededor del 70%.

Tabla 10 Pruebas chi cuadrado satisfacción vs modo y tempo.

Variable dependiente	Medida	P1		P2		P3		P4		P5	
		Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)
Tempo	Chi-cuadrado de Pearson	10,590	,001	18,065	,000	10,581	,001	16,817	,000	13,607	,000
	Corrección por continuidadb	9,689	,002	16,881	,000	9,681	,002	15,637	,000	12,581	,000
	Razón de verosimilitudes	10,685	,001	18,352	,000	10,676	,001	17,138	,000	13,770	,000
	Estadístico exacto de Fisher										
	Asociación lineal por lineal	10,537	,001	17,975	,000	10,528	,001	16,733	,000	13,539	,000
	N de casos válidos	200		200		200		200		200	
Modo	Chi-cuadrado de Pearson	37,013	,000	6,503	,011	19,222	,000	5,491	,019	20,612	,000
	Corrección por continuidadb	35,312	,000	5,801	,016	18,002	,000	4,826	,028	19,344	,000
	Razón de verosimilitudes	38,254	,000	6,540	,011	19,542	,000	5,524	,019	20,992	,000
	Estadístico exacto de Fisher										
	Asociación lineal por lineal	36,828	,000	6,471	,011	19,126	,000	5,464	,019	20,509	,000
	N de casos válidos	200		200		200		200		200	

Como se mencionaba anteriormente, las pruebas Chi Cuadrado permitirán entender si existe relación de dependencia o no entre las variables explicativas y la variable respuesta, que en este caso es la de Satisfacción Dicotómica.

Se observa entonces, que con un nivel de confianza del 95%, en términos generales todas las piezas musicales resultan significativas, esto teniendo en cuenta que el p valor en todos los casos es menor a 0,05, lo que implica que en mayor o menor medida, dependiendo del caso, siempre se rechaza la hipótesis nula H_0 de que No hay independencia entre las variables explicativas Modo y Tempo Musical y la variable dependiente Satisfacción dicotómica, es decir que si existe dependencia entre las mismas, lo que en términos de lo que se espera genera una primera luz respecto a una posible relación causal entre dichas variables.

Tabla 11 *Medidas de dirección e intensidad. Satisfacción vs modo y tempo.*

Var. Dependiente	Medidas	P1			P2			P3			P4			P5				
		Valor	Error tip. asint.a	T aproxb	Sig. Aprox	Valor	Error tip. asint.a	T aproxb	Sig. Aprox	Valor	Error tip. asint.a	T aproxb	Sig. Aprox	Valor	Error tip. asint.a	T aproxb	Sig. Aprox	
Tempo	Nominal por	Phi	,230		,001	,301		,000	,230		,001	,290		,000	,261		,000	
	nominal	V de Cramer	,230		,001	,301		,000	,230		,001	,290		,000	,261		,000	
	intervalo	C. Contingencia	,224		,001	,288		,000	,224		,001	,279		,000	,252		,000	
	Ordinal por	R de Pearson	,230	,069	3,327	,001	,301	,067	4,434	,000	,230	,069	3,326	,001	,290	,067	4,263	,000
	ordinal	Correlación de Spearman	,230	,069	3,327	,001	,301	,067	4,434	,000	,230	,069	3,326	,001	,290	,067	4,263	,000
	N de casos válidos		200			200			200			200			200			200
Modo	Nominal por	Phi	,430		,000	,180		,011	,310		,000	,166		,019	,321		,000	
	nominal	V de Cramer	,430		,000	,180		,011	,310		,000	,166		,019	,321		,000	
	intervalo	C. Contingencia	,395		,000	,177		,011	,296		,000	,163		,019	,306		,000	
	Ordinal por	R de Pearson	,430	,064	6,706	,000	,180	,070	2,580	,011	,310	,067	4,588	,000	,166	,070	2,364	,019
	ordinal	Correlación de Spearman	,430	,064	6,706	,000	,180	,070	2,580	,011	,310	,067	4,588	,000	,166	,070	2,364	,019
	N de casos válidos		200			200			200			200			200			200

Como ya se dijo anteriormente, el paso a seguir después de identificar que existe la relación entre las variables explicativas y las dependientes, es ver que intensidad tiene y su dirección, así con la ayuda de la Tabla 11, se puede ver que está presente una intensidad que puede considerarse media, de la relación entre las variables Modo y Tempo Musical con la de Satisfacción Dicotómica, tal y como lo muestran los diferentes coeficientes que en general se usan para medir esto. De otro lado se observa significancia en todos los p valores de los mencionados coeficientes ya que se encuentran por debajo de 0,05, reforzando el rechazo de la hipótesis nula de que existe independencia entre las variables Modo y Tempo versus Satisfacción.

Finalmente, respecto a las piezas musicales se aprecia una intensidad relativamente mayor de la variable Modo en las piezas 1 y 5, mientras que en la variable Tempo musical los valores de intensidad de dichos coeficientes se mantienen más estables entre las diferentes piezas musicales en general, solo con valores más altos en la pieza 2.

Análisis bivariado variable Número de Parpadeos.

Al igual que con la variable de Satisfacción, a continuación, se realizará el análisis del respectivo cruce de las variables Modo y Tempo Musical con la variable dependiente Número

de Parpadeos Dicotómica, en donde la tabla de contingencia resumida será de utilidad para ver si hay algún patrón de concentración marcado y consistente en los casos para las diferentes piezas musicales y la comprobación estadística de la existencia o no de relaciones de dependencia entre la variables analizadas y la dirección e intensidad de la misma.

Tabla 12 *Tablas de contingencia número de parpadeos vs modo y tempo.*

Tempo															
Categorías	P1			P2			P3			P4			P5		
	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total
Pocos	63%	29%	46%	76%	38%	57%	87%	34%	61%	82%	43%	63%	84%	47%	66%
Hartos	37%	71%	54%	24%	62%	43%	13%	66%	40%	18%	57%	38%	16%	53%	35%
Modo															
Categorías	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total
	Pocos	50%	42%	46%	66%	48%	57%	73%	48%	61%	71%	54%	63%	76%	55%
Hartos	50%	58%	54%	34%	52%	43%	27%	52%	40%	29%	46%	38%	24%	45%	35%

De acuerdo a lo que se puede apreciar en la Tabla 12, es evidente que al igual que en el caso anterior, las variables independientes Tempo y Modo ejercen influencia sobre los resultados de la variable dependiente Número de Parpadeos dicotómica, ya que en la mayoría de los casos, se observan concentraciones mayores al 50% para los cruces de las categorías Hartos-Rápido y Pocos-Lento en la variable Independiente Tempo VS Número de Parpadeos dicotómica y en los cruces de las categorías Pocos-Menor y Hartos-Mayor para la variable independiente Modo Musical VS Número de parpadeos Dicotómica.

De otro lado, se observa una mayor influencia de la variable Tempo Musical en comparación con la de Modo, ya que las concentraciones de los casos en la variable dependiente según las categorías Rápido-Muchos y Lento-Pocos son en su mayoría mayores al 60%.

Tabla 13 *Pruebas chi cuadrado número de parpadeos vs modo y tiempo.*

Variable dependiente	Medida	P1		P2		P3		P4		P5	
		Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)
Tempo	Chi-cuadrado de Pearson	10,590	,001	18,065	,000	10,581	,001	16,817	,000	13,607	,000
	Corrección por continuidad ^b	9,689	,002	16,881	,000	9,681	,002	15,637	,000	12,581	,000
	Razón de verosimilitudes	10,685	,001	18,352	,000	10,676	,001	17,138	,000	13,770	,000
	Estadístico exacto de Fisher										
	Asociación lineal por lineal	10,537	,001	17,975	,000	10,528	,001	16,733	,000	13,539	,000
	N de casos válidos	200		200		200		200		200	
Modo	Chi-cuadrado de Pearson	37,013	,000	6,503	,011	19,222	,000	5,491	,019	20,612	,000
	Corrección por continuidad ^b	35,312	,000	5,801	,016	18,002	,000	4,826	,028	19,344	,000
	Razón de verosimilitudes	38,254	,000	6,540	,011	19,542	,000	5,524	,019	20,992	,000
	Estadístico exacto de Fisher										
	Asociación lineal por lineal	36,828	,000	6,471	,011	19,126	,000	5,464	,019	20,509	,000
	N de casos válidos	200		200		200		200		200	

Las pruebas de independencia Chi Cuadrado entre la variable explicada Número de parpadeos Dicotómica y Las Variables Dependientes Modo y Tempo Musical, son significativas ya que en todas las 5 piezas musicales el p valor se encuentra siempre por debajo de 0,05, indicando así la existencia de evidencia estadística de que las variables explicativas no son independientes de la variable respuesta analizada ya que siempre se rechaza esta hipótesis nula de independencia.

Tabla 14 *Medidas de dirección e intensidad. número de parpadeos vs modo y tiempo.*

Var. Dependiente	Medidas	P1					P2					P3					P4					P5							
		Valor	Error tip.	T	aprox. b	Sig. Aprox	Valor	Error tip.	T	aprox. b	Sig. Aprox	Valor	Error tip.	T	aprox. b	Sig. Aprox	Valor	Error tip.	T	aprox. b	Sig. Aprox	Valor	Error tip.	T	aprox. b	Sig. Aprox			
Tempo	Nominal por Phi	,341			,000	,384			,000	,542			,000	,403			,000	,389			,000	,389			,000	,389			,000
	Intervalo por V de Cramer	,341			,000	,384			,000	,542			,000	,403			,000	,389			,000	,389			,000	,389			,000
	Ordinal por C. Contingencia	,323			,000	,358			,000	,477			,000	,374			,000	,363			,000	,363			,000	,363			,000
	N de casos válidos	,341	,066	5,106	,000	,384	,065	5,848	,000	,542	,058	9,077	,000	,403	,063	6,192	,000	,389	,063	5,945	,000	,389	,063	5,945	,000	,389	,063	5,945	,000
	Intervalo por R de Pearson	,341	,066	5,106	,000	,384	,065	5,848	,000	,542	,058	9,077	,000	,403	,063	6,192	,000	,389	,063	5,945	,000	,389	,063	5,945	,000	,389	,063	5,945	,000
	Ordinal por Correlación de	,341	,066	5,106	,000	,384	,065	5,848	,000	,542	,058	9,077	,000	,403	,063	6,192	,000	,389	,063	5,945	,000	,389	,063	5,945	,000	,389	,063	5,945	,000
Modo	Nominal por Phi	,080			,256	,182			,010	,256			,000	,176			,013	,221			,002	,221			,002	,221			,002
	Intervalo por V de Cramer	,080			,256	,182			,010	,256			,000	,176			,013	,221			,002	,221			,002	,221			,002
	Ordinal por C. Contingencia	,080			,256	,179			,010	,248			,000	,173			,013	,216			,002	,216			,002	,216			,002
	N de casos válidos	,080	,070	1,133	,259	,182	,069	2,601	,010	,256	,068	3,722	,000	,176	,069	2,510	,013	,221	,068	3,187	,002	,221	,068	3,187	,002	,221	,068	3,187	,002
	Intervalo por R de Pearson	,080	,070	1,133	,259	,182	,069	2,601	,010	,256	,068	3,722	,000	,176	,069	2,510	,013	,221	,068	3,187	,002	,221	,068	3,187	,002	,221	,068	3,187	,002
	Ordinal por Correlación de	,080	,070	1,133	,259	,182	,069	2,601	,010	,256	,068	3,722	,000	,176	,069	2,510	,013	,221	,068	3,187	,002	,221	,068	3,187	,002	,221	,068	3,187	,002

Según lo observado en la tabla 14, tanto la variable Tempo como la variable Modo presentan asociación positiva con la dependiente Número de Parpadeos Dicotómica, esto teniendo en cuenta que todas las medidas tienen valores positivos en todos los casos, sin embargo, se observa que la significancia es más robusta en todos los casos para la variable

Tempo ya que siempre está por debajo de 0,05 y además siempre muy cercana a 0, mientras que para la variable Modo esto no sucede ya que, aunque en la mayoría de los casos la significancia está por debajo de 0,05, no está tan cercana a 0 como en la variable Tempo, también se aprecia que en la pieza 1 incluso el resultado no es significativo en ninguna de las medidas de asociación teniendo en cuenta que los valores p están sobre el mencionado 0,05. En general la variable Tempo presenta grados de asociación relativamente mayores respecto a la variable Modo, valores que en general están sobre 0,3 y que en los casos de las piezas 3 y 4 llegan a ser superiores a 0,5 y 0,4 respectivamente en la mayoría de los casos.

Análisis bivariado variable Tiempo de Respuesta.

A continuación se mostrará el análisis cruzado entre la variable dependiente dicotómica Tiempo de Respuesta y las variables explicativas Modo y Tempo Musical, para de igual forma que en los casos anteriores observar si existe algún tipo de relación de asociación y la dirección y fuerza de la misma en caso de la existencia.

Tabla 15 *Tablas de contingencia tiempo de respuesta vs modo y tempo.*

Tempo															
Categorías	P1			P2			P3			P4			P5		
	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total
Lento	74%	45%	60%	62%	48%	55%	72%	44%	58%	71%	43%	57%	66%	46%	56%
Rápido	26%	55%	41%	38%	52%	45%	28%	56%	42%	29%	57%	43%	34%	54%	44%
Modo															
Categorías	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total
	Lento	67%	52%	60%	69%	41%	55%	71%	45%	58%	72%	42%	57%	65%	47%
Rápido	33%	48%	41%	31%	59%	45%	29%	55%	42%	28%	58%	43%	35%	53%	44%

De acuerdo a lo que se observa en la Tabla 15 parece existir relación entre las categorías de las variables dependientes frente a esta variable explicada, ya que se observa una distribución mayor siempre al 50% en las combinaciones de las categorías Lento-Lento y Rápido-Rápido lo que sugiere un patrón de concentración. De otro lado, no es evidente ninguna diferencia a primera vista en los efectos que pueda causar la variable Tempo y la Variable Modo sobre la variable dependiente.

Tabla 16 Pruebas chi cuadrado tiempo de respuesta vs modo y tiempo.

Variable dependiente	Medida	P1		P2		P3		P4		P5	
		Valor	asintótica (bilateral)	Valor	asintótica (bilateral)	Valor	asintótica (bilateral)	Valor	asintótica (bilateral)	Valor	asintótica (bilateral)
Tempo	Chi-cuadrado de Pearson	17,450	,000	3,960	,047	16,092	,000	15,993	,000	8,117	,004
	Corrección por continuidad	16,267	,000	3,414	,065	14,963	,000	14,871	,000	7,325	,007
	Razón de verosimilitudes	17,756	,000	3,973	,046	16,340	,000	16,233	,000	8,176	,004
	Estadístico exacto de Fisher										
	Asociación lineal por lineal	17,363	,000	3,940	,047	16,011	,000	15,914	,000	8,076	,004
	N de casos válidos	200		200		200		200		200	
Modo	Chi-cuadrado de Pearson	4,669	,031	15,838	,000	13,875	,000	18,360	,000	6,575	,010
	Corrección por continuidad	4,067	,044	14,727	,000	12,828	,000	17,156	,000	5,864	,015
	Razón de verosimilitudes	4,690	,030	16,064	,000	14,059	,000	18,677	,000	6,613	,010
	Estadístico exacto de Fisher										
	Asociación lineal por lineal	4,645	,031	15,759	,000	13,806	,000	18,268	,000	6,542	,011
	N de casos válidos	200		200		200		200		200	

Como se puede observar la Tabla 16 confirma que en general estadísticamente hay evidencia respecto a la existencia de relación de dependencia entre las variables Tempo y Modo y Tiempo de Respuesta Dicotómico al estar todos los valores p en su mayoría por debajo de 0,05.

Tabla 17 Medidas de dirección e intensidad. Número de parpadeos vs modo y tiempo.

Var. Dependiente	Medidas	P1				P2				P3				P4				P5				
		Valor	Error tip.	T	Sig. Aprox	Valor	Error tip.	T	Sig. Aprox	Valor	Error tip.	T	Sig. Aprox	Valor	Error tip.	T	Sig. Aprox	Valor	Error tip.	T	Sig. Aprox	
Tempo	Nominal por nominal	Phi	,295		,000	,141			,047	,284			,000	,283			,000	,201			,004	
		V de Cramer	,295		,000	,141			,047	,284			,000	,283			,000	,201			,004	
		C. Contingencia	,283			,139			,047	,273			,000	,272			,000	,197			,004	
	Intervalo por ordinal	R de Pearson	,295	,067	4,350	,000	,141	,070	2,000	,047	,284	,068	4,162	,000	,283	,068	4,148	,000	,201	,069	2,894	,004
		Correlación de	,295	,067	4,350	,000	,141	,070	2,000	,047	,284	,068	4,162	,000	,283	,068	4,148	,000	,201	,069	2,894	,004
		N de casos válidos	200				200				200				200				200			
Modo	Nominal por nominal	Phi	,153		,031	,281			,000	,263			,000	,303			,000	,181			,010	
		V de Cramer	,153		,031	,281			,000	,263			,000	,303			,000	,181			,010	
		C. Contingencia	,151		,031	,271			,000	,255			,000	,290			,000	,178			,010	
	Intervalo por ordinal	R de Pearson	,153	,070	2,175	,031	,281	,068	4,127	,000	,263	,068	3,842	,000	,303	,067	4,474	,000	,181	,069	2,594	,010
		Correlación de	,153	,070	2,175	,031	,281	,068	4,127	,000	,263	,068	3,842	,000	,303	,067	4,474	,000	,181	,069	2,594	,010
		N de casos válidos	200				200				200				200				200			

Finalmente, en la Tabla 17 se muestra la dirección e intensidad de las relaciones que se encontraron en el punto anterior, como se puede observar en todas las piezas musicales la relación es positiva, sin embargo, no en todas las piezas musicales la relación tiene el mismo grado de intensidad ya que la 1 3 y 4, presentan valores cercanos a 0,3 en la variable Tempo, mientras que la 2 y la 5 presentan valores más bajos. De otro lado, respecto a la variable Modo los valores más altos cercanos a 0,3 están en las piezas 2,3 y4, mientras que en las piezas 1 y 5 los valores de los estadísticos de intensidad de la dependencia son menores.

Análisis bivariado variable Latencia.

A continuación, se mostrará el análisis cruzado entre la variable dependiente dicotómica Latencia y las variables explicativas Modo y Tempo Musical con el objetivo también verificar la existencia de dependencia y su intensidad y dirección en los casos en los que aplique.

Tabla 18 *Tablas de contingencia latencia vs modo y tempo.*

Tempo															
Categorías	P1			P2			P3			P4			P5		
	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total	Lento	Rápido	Total
Poca	75%	50%	63%	70%	50%	60%	72%	52%	62%	71%	49%	60%	68%	49%	59%
Harta	25%	50%	38%	30%	50%	40%	28%	48%	38%	29%	51%	40%	32%	51%	42%
Modo															
Categorías	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total
	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total	Menor	Mayor	Total
Poca	74%	51%	63%	71%	49%	60%	73%	51%	62%	73%	47%	60%	66%	51%	59%
Harta	26%	49%	38%	29%	51%	40%	27%	49%	38%	27%	53%	40%	34%	49%	42%

La Tabla 18 evidencia un patrón de concentraciones diferente al de los casos anteriores en el sentido en que para esta variable dependiente las concentraciones de las combinaciones que representan las casillas superior izquierda de cada pieza musical para cada variable independiente concentran en su mayoría más de un 70% en todos los casos con excepción de la pieza 5, lo que significa que no es tan clara, a primera vista, la relación que se espera en donde menor latencia se relacione con Tempo Rápido y Modo Mayor.

Tabla 19 *Pruebas chi cuadrado latencia vs modo y tempo.*

Variable dependiente	Medida	P1		P2		P3		P4		P5	
		Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)	Valor	Sig. asintótica (bilateral)
Tempo	Chi-cuadrado de Pearson	13,333	,000	8,333	,004	8,489	,004	10,083	,001	7,435	,006
	Corrección por continuidad ^b	12,288	,000	7,521	,006	7,661	,006	9,188	,002	6,673	,010
	Razón de verosimilitudes	13,529	,000	8,402	,004	8,566	,003	10,185	,001	7,487	,006
	Estadístico exacto de Fisher										
	Asociación lineal por lineal	13,267	,000	8,292	,004	8,447	,004	10,033	,002	7,398	,007
	N de casos válidos	200		200		200		200		200	
Modo	Chi-cuadrado de Pearson	11,285	,001	10,083	,001	10,272	,001	14,083	,000	4,634	,031
	Corrección por continuidad ^b	10,325	,001	9,188	,002	9,359	,002	13,021	,000	4,037	,045
	Razón de verosimilitudes	11,424	,001	10,185	,001	10,384	,001	14,284	,000	4,654	,031
	Estadístico exacto de Fisher										
	Asociación lineal por lineal	11,229	,001	10,033	,002	10,220	,001	14,013	,000	4,611	,032
	N de casos válidos	200		200		200		200		200	

La Tabla 19 evidencia que existe relación de dependencia entre las variables Modo y Tempo y la variable de Latencia, tal y como se observa los valores p para las 2 variables en todas las piezas musicales resultan significativos para las respectivas pruebas chi cuadrado en la medida en que están por debajo de 0,05.

Tabla 20 *Medidas de dirección e intensidad latencia vs modo y tempo.*

Var. Dependiente	Medidas	P1				P2				P3				P4				P5				
		Valor	Error tip.	T	aprox Sig. Aprox	Valor	Error tip.	T	aprox Sig. Aprox	Valor	Error tip.	T	aprox Sig. Aprox	Valor	Error tip.	T	aprox Sig. Aprox	Valor	Error tip.	T	aprox Sig. Aprox	
Tempo	Nominal por nominal	Phi	,258			,000	,204			,004	,206			,004	,225			,001	,193			,006
		V de Cramer	,258			,000	,204			,004	,206			,004	,225			,001	,193			,006
		C.	,250			,000	,200			,004	,202			,004	,219			,001	,189			,006
	Intervalo por Ordinal	R de	,258	,068	3,761	,000	,204	,069	2,934	,004	,206	,069	2,963	,003	,225	,069	3,242	,001	,193	,069	2,765	,006
	Ordinal por Ordinal	Correlación	,258	,068	3,761	,000	,204	,069	2,934	,004	,206	,069	2,963	,003	,225	,069	3,242	,001	,193	,069	2,765	,006
	N de casos válidos	200				200				200				200				200				
Modo	Nominal por nominal	Phi	,238			,001	,225			,001	,227			,001	,265			,000	,152			,031
		V de Cramer	,238			,001	,225			,001	,227			,001	,265			,000	,152			,031
		C.	,231			,001	,219			,001	,221			,001	,256			,000	,150			,031
	Intervalo por Ordinal	R de	,238	,068	3,441	,001	,225	,069	3,242	,001	,227	,069	3,274	,001	,265	,068	3,873	,000	,152	,070	2,167	,031
	Ordinal por Ordinal	Correlación	,238	,068	3,441	,001	,225	,069	3,242	,001	,227	,069	3,274	,001	,265	,068	3,873	,000	,152	,070	2,167	,031
	N de casos válidos	200				200				200				200				200				

En la tabla 20 se aprecian relaciones positivas en todas las piezas musicales que se pueden categorizar en una intensidad media, sin embargo, como es de esperarse no en todas las piezas musicales los indicadores muestran la misma fuerza de intensidad en la relación, así, se puede ver que la pieza Número 5 arroja un indicador de intensidad menor que las demás, mientras que la 1 y la 4 muestran comportamientos opuestos en donde se muestran como las más influyentes en la variable dependiente Latencia Dicotómica, esto es así especialmente para la variable Tempo en la pieza 1 y para la variable Modo en la pieza 4.

Conclusiones Análisis bivariado.

Respecto al análisis bivariado podemos observar que en su gran mayoría existe evidencia estadística de la relación de dependencia que hay entre las variables independientes Modo y Tempo Musical con las diferentes variables explicadas, esto ya que en la gran mayoría,

de los casos como ya se ha dicho, el p valor de las pruebas Chi- Cuadrado se encuentra por debajo de 0,05, la única excepción a este comportamiento corresponde a la prueba de la variable modo versus la dependiente Número de parpadeos dicotómica para la pieza musical 1, en donde el p valor es de 0,25.

Por otro lado, también se hace evidente que no todas las piezas musicales tienen un comportamiento de influencia persistente sobre las diferentes variables dependientes analizadas, sin embargo las piezas 1 y 4 se diferencian de las demás en la medida en que si presentan dicho comportamiento, especialmente esta última con intensidades relativamente mayores que las demás piezas particularmente en la variable independiente Modo Musical.

En las variables de Satisfacción Dicotómica y Tiempo de Respuesta parece darse la relación esperada en la medida en que el Tempo Rápido y Modo Mayor generan mayores respuestas en los valores 1 de la variable de Satisfacción dicotómica (“Satisfecho”), de la misma forma ocurre con la variable Tiempo de Respuesta en donde Tiempos Rápidos que significarían satisfacción están influenciados por los valores 1 de las respectivas variables independientes (Rápido para Tempo y Mayor para Modo Musical), sin embargo no en todos los casos ocurre esto, teniendo en cuenta que la variable Latencia no presenta en el análisis el comportamiento esperado de Satisfacción, ya que las concentraciones de casos que generan Tempo Rápido y Modo Mayor en Poca latencia (Satisfacción) no son mayores al 50% en todos los casos como se esperaría, de la misma forma ocurre con los casos que se esperaría se concentraran más en insatisfacción (Latencia larga) versus Modo Menor y Tempo Lento.

Finalmente la variable Número de parpadeos parece mostrar una relación entre Tempo Lento y pocos Parpadeos y Rápido con Hartos y entre modos menores y Pocos parpadeos y Modos Mayores y Hartos parpadeos. Aunque la variable Número de parpadeos se introduce en el análisis sin ningún comportamiento esperado inicial, pudiera estarse relacionando de la

misma forma con la satisfacción teniendo en cuenta el comportamiento que se observa en las variables independientes y que es similar al de la variable Satisfacción dicotómica.

Conclusiones Generales Análisis Exploratorio

Como ya se mencionaba anteriormente, el análisis exploratorio es importante en la medida en que empieza a dilucidar los comportamientos y relaciones que existen en las variables independientes versus las dependientes y permite evidenciar si finalmente se dan los comportamientos esperados por el investigador.

Una primera conclusión importante que arroja el análisis univariado, como se vio anteriormente en el respectivo apartado, es el hecho de que las variables dependientes mostraran comportamientos de agrupación definidos en las variables dependientes originales, esto permite evidenciar que en un principio existe algún tipo de efecto discriminativo en los resultados de dichas variables lo que en efecto facilita la creación de las variables dicotómicas que se usarán en las técnicas posteriormente aplicadas.

De otro lado también se observa que las curvas tienden a concentrarse en general en valores bajos de las respectivas escalas de cada una de las variables independientes analizadas, lo que dependiendo de la variable explicada que se vea, podría significar tendencias de satisfacción definidas en estas variables y las respectivas piezas musicales. Sin embargo, dependiendo la variable explicada se observan comportamientos que pueden ser o no consistentes en todas las piezas musicales, por ejemplo en las variables dicotómicas Tiempo de Respuesta y Latencia todas las piezas musicales concentran sus respuestas en los valores bajos de la escala (Rápido y Poca), mientras que en las variables de Satisfacción dicotómica y Número de parpadeos la pieza 1 se comporta diferente a las demás, así, lo anterior sugiere la influencia de la pieza musical en estas variables donde los comportamientos si difieren entre piezas. Se observa que en general las piezas 1, 4 y 5 muestran comportamientos consistentes y

sobresalientes de las demás, lo que en una primera vista puede sugerir que estas piezas influyen en la satisfacción de los individuos de forma más marcada que las demás.

Respecto al análisis bivariado es importante ver que las pruebas de dependencia Chi cuadrado, para todas las variables explicadas versus las independientes dan significativas en la gran mayoría de los casos, solo se presenta una excepción en la pieza 1 en la prueba para la variable Modo vs Número de parpadeos dicotómica, de resto como ya se ha dicho la evidencia estadística sugiere la existencia de relación de dependencia entre las variables explicadas y las independientes lo que se convierte en un resultado importante para proceder con los posteriores análisis partiendo de la base de que efectivamente existe alguna relación entre estos dos tipos de variables.

También resulta interesante y de igual importancia el hecho de que no solo existe relación de dependencia entre estas variables, sino que en muchos de los casos, de acuerdo a los estadísticos que miden la intensidad de las relaciones (V de Cramer, el Coeficiente de Contingencia, Phi, entre otros), la fuerza de la misma es importante al encontrarse en general en rangos de entre el 0,2 y llegar en algunos casos hasta niveles del 0,5 como en la pieza 3 para la variable Número de Parpadeos dicotómica.

Finalmente, respecto a las variables independientes se observa que en general no hay un comportamiento persistente en alguna de ellas que permita concluir que su influencia sobre los resultados en la variable dependiente respectiva sea determinante, solo se encuentra una excepción a esto y es en la variable Número de Parpadeos Dicotómica en la que se observan indicadores de intensidad mayores en todas las piezas para la variable Tempo respecto a los observados en la variable Modo.

Perspectiva No Paramétrica Del Análisis

Debido al alcance del trabajo es necesario tener en cuenta cuales son las técnicas estadísticas que se utilizan con más frecuencia en la validación de los resultados obtenidos en

los experimentos con aplicación a la psicología, para esto es importante distinguir entre los métodos estadísticos paramétricos y los no paramétricos ya que según las características de los datos del análisis unos métodos pueden ajustarse mejor que otros.

Los métodos paramétricos hacen referencia directa a los parámetros de la población y sobre ellos realizan sus supuestos y estimaciones, en tanto que los supuestos de los métodos no paramétricos, por el contrario, no hacen referencia alguna a los parámetros poblacionales ya que en muchos de los casos no se tiene información sobre la distribución de los datos originales es así que al no tener parámetros la alternativa de análisis debe enfocarse más en comparaciones y análisis sobre la distribución de los datos como tal, a estas técnicas, de acuerdo a autores como Villazon y Brenard (1992), se les suele llamar con frecuencia pruebas de rango o de orden .

En el caso de datos que son nominales y ordinales, en los que el cumplimiento de los requerimientos que exigen los métodos paramétricos tradicionales, son difíciles de alcanzar, el uso de los métodos no paramétricos se convierte en una alternativa práctica para poder trabajar dichos datos estadísticamente evitando la dificultad que suponen las mencionadas técnicas paramétricas (Monroy 2009). En estudios de Mercado y aplicados a la psicología del consumidor resulta muy frecuente la obtención de este tipo de datos, tal y como sucede en el caso del presente análisis, por lo que puede resultar bastante útil la aplicación de la técnica no paramétrica que mejor se ajuste para la obtención de conclusiones válidas y apropiadas respecto a los efectos que tienen las variables independientes seleccionadas frente a cada una de las variables respuesta de satisfacción.

Podemos encontrar una clasificación de las pruebas no paramétricas según el número de muestras que se analicen y si estas se encuentran relacionadas o no, criterio que se utiliza en la selección de la prueba que se aplicará en el presente análisis dada la naturaleza de los datos utilizados:

Pruebas para una muestra

- Prueba de chi cuadrado de Pearson
- Prueba Binomial
- Prueba de rachas
- Prueba de Kolmogorov Smirnov (K-S)

Pruebas para dos muestras relacionadas

- Prueba de Mc Nemar
- Prueba de los signos
- Prueba de Wilcoxon

Pruebas K muestras relacionadas

- Prueba de Friedman
- Prueba de Cochran
- Coeficiente de concordancia de W de kendall

Como se puede observar, claramente el análisis a realizar se enmarca en el campo de dos muestras relacionadas ya que los datos tanto de las variables respuesta como de las independientes es extraída de la misma muestra de individuos, y se quiere observar la distribución de los datos para cada variable independiente según los resultados obtenidos en cada variable respuesta, de este modo se selecciona la prueba de Wilcoxon como la más apropiada para aplicar en este análisis.

Aplicación de Pruebas de Rangos De Wilcoxon

En este aparte se abordará el análisis de los datos por medio de una de las pruebas estadísticas no paramétricas más usadas en estas situaciones, en las que se busca inferir sobre los efectos que tienen los niveles de las variables independientes dicotómicas, Tempo y Modo Musical, sobre los valores que toman las variables respuesta estudiadas relacionadas con la satisfacción. De este modo, se aplicarán pruebas de signo de rangos de Wilcoxon a cada una

de las variables independientes sobre cada variable respuesta para las 5 piezas musicales con el objetivo de generar conclusiones respecto al poder de influencia que tienen dichos niveles de variación en el Tempo y Modo Musical sobre la satisfacción y en cuál de las piezas musicales se observa mejor este efecto en el caso en que efectivamente se presente.

Para la aplicación de las pruebas se toma como referencia el modelo psicológico propuesto por Scherer y Zetner (2001) para el análisis de los efectos que produce la música sobre las emociones en los seres humanos. El presente trabajo solo pretende identificar evidencia de la influencia que se espera observar sobre la satisfacción, sin embargo se puede mencionar la utilidad que este tipo de estudios tiene en la práctica ya que por medio del entendimiento de estos efectos y sus causas, se podrían generar herramientas prácticas en campos como la publicidad y el mercadeo en donde se logren crear asociaciones de productos con efectos positivos de satisfacción a través de la creación de piezas musicales o sonidos que incluyan características que produzcan estos efectos en el consumidor.

El modelo que plantean estos autores consta de 5 elementos que se agrupan en 2 grandes categorías a saber: características estructurales y no estructurales. En el primer grupo se encuentran las llamadas características de Segmento que son más elementos puramente acústicos y que están sujetos con mayor facilidad a análisis cuantitativos, podemos encontrar así dentro de estos elementos los tonos, los acordes y los intervalos musicales. De otro lado tenemos las características de Supersegmento con mayor complejidad que las anteriores teniendo en cuenta que corresponden articulaciones en el tiempo de las mismas, están en este grupo la Melodía, el Tempo, el Ritmo y la Armonía. En el otro gran grupo de características no estructurales están todas aquellas que no tienen que ver directamente con la música sino más con el contexto, con el oyente mismo y con el músico que interpreta la pieza.

La intención de la prueba piloto realizada con la que fueron recolectados los datos se concentra en las características del primer grupo y neutraliza de la mejor forma las del segundo

controlando: el ambiente al ser realizada en un espacio aislado de sonidos y perturbaciones externas y piezas musicales neutras que solo tienen variaciones de tempo y modo musical y ninguna relación con alguna canción o melodía que puedan conocer los participantes y que pudiera generar algún sesgo en las medidas tomadas de satisfacción.

De este modo se aplican las pruebas a cada una de las 5 piezas y se observan las diferencias en distribución que causan los niveles de las variables Tempo y Modo Musical (cada una por separado) respecto a los valores de las variables respuesta estudiadas (Satisfacción, Tiempo de Respuesta, Duración del Parpadeo, Número de Parpadeos por minuto).

De este Modo las poblaciones que se comparan corresponden a los valores de cada variable respuesta definidos para el nivel dicotómico que toma la variable independiente que se está analizando. Así, por ejemplo se contrastan los valores que toma la variable Satisfacción para el modo Mayor versus los valores que toma esta variable en el Modo Menor, se tiene entonces la aplicación de las pruebas para las variables respuestas y las siguientes poblaciones que definen las parejas del contraste según sus niveles, tal y como se observa en la siguiente Tabla:

Tabla 21 *Combinaciones para aplicación de pruebas de Wilcoxon*

Variable Respuesta	Modo	Tempo
Satisfacción	Mayor Vs Menor	Rápido Vs Lento
Tiempo de Respuesta		
Duración del Parpadeo		
Numero de Parpadeos X Minuto		

Son en resumen 8 los contrastes que se realizan para cada una de las piezas musicales, ya que cada variable independiente se contrasta con cada una de las 4 dependientes, corriendo el análisis tenemos los siguientes estadísticos de contraste en las 5 piezas musicales, tal y como se ve en la siguiente tabla:

Tabla 22 *Aplicación prueba signos de rangos de Wilcoxon, valores p*

Pieza Musical	Modo (Menor y Mayor Vs Variable)				Tempo (Rápido y Lento Vs Variable)			
	Satisfacción	Tiempo Respuesta	Duración Parpadeo	Parpadeo por minuto	Satisfacción	Tiempo Respuesta	Duración Parpadeo	Parpadeo por minuto
1	.000	.001	.002	.000	.000	.000	.001	.000
2	.000	.001	.002	.000	.000	.000	.001	.000
3	.024	.000	.012	.000	.000	.008	.001	.000
4	.000	.000	.001	.000	.026	.000	.003	.000
5	.024	.000	.025	.001	.000	.000	.045	.000

Se puede observar entonces que en términos generales las pruebas para las 5 piezas musicales arrojan resultados significativos respecto al p valor que siempre es menor a 0,05 (al realizar la prueba con un 95% de confianza) indicando así que se rechaza en todos los casos la hipótesis nula de que no hay diferencias en los datos de cada una de las variables respuesta para cada variable independiente por aparte en sus 2 niveles. Esto indica que la distribución de los datos que conforman los valores de cada variable respuesta es diferente al aplicar los niveles de las variables independientes lo que puede traducirse a la evidencia estadística de que existe un efecto causal de las variables independientes sobre las de satisfacción.

Sin embargo, podemos ver también en los resultados que en el caso de las piezas musicales 3 y 5 estos efectos no se presentan de forma tan fuerte para las variables de Satisfacción y número de parpadeos en el Modo Mayor como en las demás piezas musicales tal y como lo muestran valores p más lejanos a 0 respecto a las demás variables respuesta en estas dos piezas musicales.

También se presenta un resultado similar en el contraste de la variable independiente Tempo versus la independiente Duración del Parpadeo en donde se observa un p valor de 0.045, muy cercano al 0,05, lo que significa un efecto no tan fuerte como se menciona en el párrafo anterior.

Perspectiva Paramétrica del Análisis

Desde un punto de vista paramétrico, se puede observar que se tiene un caso de variables cualitativas independientes o explicativas, que como ya se ha mencionado anteriormente, corresponden a variables tempo y modo musical. Se busca entonces encontrar efectos de dichas variables sobre las de satisfacción que son variables respuesta que pueden usarse de modo binario en términos de si la pieza musical fue calificada como satisfactoria o no. Desde este punto de vista se ajusta muy bien la técnica de regresión logística según la estructura de los datos en donde se trabajan variables binarias explicativas y respuesta, también puede analizarse el problema mediante la técnica de árboles de clasificación que como se verá más adelante es una técnica muy usada para predecir y clasificar eventos binarios al igual que la regresión logística, pero mediante reglas de clasificación definidas por la probabilidad de que sucedan dichos eventos.

Modelos de Elección y Preferencias

Estos modelos son los más utilizados para estudios en las preferencias de los consumidores, están muy relacionados con la teoría de las preferencias declaradas (Mazzochi 2008), teorías que son muy usadas en estudios del consumidor y estudios económicos para conocer la reacción que tendrían las personas ante nuevas opciones, se pueden encontrar dentro de dichos estudios los de preferencias en encuestas (elecciones entre diferentes marcas, productos y características, diferentes tiendas, etc.), se caracterizan también por tener la posibilidad de dar a escoger entre opciones que pueden ser tanto hipotéticas como reales, estos modelos se enfocan en encontrar los determinantes de las preferencias declaradas del consumidor a partir de estas mismas.

Una de las alternativas que hay ante las teorías de elección declarada son las de elecciones reveladas que contrariamente a las primeras buscan encontrar alguna preferencia del consumidor ante su elección pero de modo indirecto, esto es, sin preguntar directamente al

consumidor sobre sus preferencias sino, en lugar de ello, recolectando información que indirectamente logre reunir lo necesario para deducir la elección final del consumidor a través de otros comportamientos, así se puede tomar el ejemplo de Mazzochi (2008) en el que se dobla el precio de un jabón para observar el comportamiento que toma el consumidor frente a esta situación, si el consumidor decide no comprar el jabón será posible inferir esta decisión con una simple inspección de su carrito de mercado?, o si el cliente compró otro jabón con el mismo precio al que antes tenía el jabón en cuestión, será esto una preferencia revelada?.

Como se puede observar en el primer grupo de teorías, de preferencias declaradas, el camino para estudiar los comportamientos del consumidor es más directo, que en las segundas ya que se parte de la decisión específica del consumidor y de lo que se trata es de identificar qué factores influyen en ella, en el segundo grupo de teorías de preferencias reveladas, se requiere un mayor esfuerzo de recolección de información para inferir el comportamiento del consumidor es por esto que en ocasiones es más sencillo y preferido el uso de las teorías de preferencia declarada.

En los modelos de elección revelada es posible utilizar regresiones con variables cuantitativas en ambos lados ya que en general se miden este tipo de variables, por ejemplo la cantidad de productos como variable dependiente de otras cuantitativas como los precios, cantidades de compras de otros productos, entre otras. Es así que un modelo de regresión lineal o multivariado con variables cuantitativas es viable. De otro lado, contrario a lo que sucede con los modelos de preferencias reveladas, en los modelos de preferencias directas o declaradas no se pueden usar este tipo de técnicas ya que son modelos construidos con variables dependientes cualitativas, su data proviene de encuestas en donde el consumidor directamente expresa su preferencia entre diferentes alternativas, las cuales conforman los valores o niveles de variables que pueden o no ser dicotómicas pero siempre categóricas, por ejemplo cuando un cliente decide o no comprar un jabón y tiene que decidir entre diferentes marcas, en este caso una

regresión estándar no es apropiada por lo que se debe utilizar alguna de las técnicas que se enmarcan dentro de los modelos de elección discreta.

Modelos de Elección Discreta

Los modelos de elección discreta generalizan los modelos de regresión a situaciones en las que la variable dependiente no es métrica o continua. Para estos casos la variable dependiente puede presentarse en varias formas: puede ser binaria (0 o 1), también puede ser ordinal tal y como sucede cuando en un cuestionario se utilizan escalas para darle diferentes opciones al encuestado en las que se reflejan varios niveles de preferencia ordenados, como: completamente en desacuerdo, en desacuerdo, de acuerdo y completamente de acuerdo, o finalmente puede ser una variable categórica que recoja un destino de vacaciones preferido de cada encuestado.

Las variables explicativas generalmente pueden ser variables métricas o variables categóricas o binarias que se pueden volver variables dummi para ser usadas dentro del análisis como se hace en el caso de una regresión tradicional; las variables dependientes que no son métricas violan los supuestos de homocedasticidad y normalidad de las regresiones lo que se soluciona con las alternativas que presentan los modelos de elección discreta.

Modelos de elección binaria.

Estos modelos se caracterizan por tener una variable dependiente “y” que puede asumir valores entre 0 y 1 de forma discreta, el modelamiento de la variable dependiente “y” en función de x se realiza a través de la explotación de una variable latente z, “y” asume valores entre 1 y 0 dependiendo de los valores que tome la variable latente z y los límites que defina un parámetro δ , la variable explicativa x está ligada a “y” por medio de la variable latente z definida por un modelo de regresión lineal de x como variable explicativa de z que a su vez define los valores de “y” dependiendo de si sobrepasa el límite definido por el parámetro δ como ya se explicó anteriormente.

Para la regresión auxiliar, es decir la que se usa para la variable latente z es necesario saber o conocer a priori como se distribuye el error, es aquí que se debe tener en cuenta el concepto de las “link functions”, o funciones de enlace las cuales especifican la relación entre z y “ y ” a través del valor esperado de la función de distribución apropiada para cada observación y_i , por ejemplo en el caso de datos binarios se podría asumir que las probabilidades de cada observación y_i siguen una distribución binomial, en este caso se pueden encontrar varias transformaciones de “ y ” que crean una variable latente z compatible con una distribución binomial

Regresión logística.

La regresión logística es una metodología que utiliza la transformación logística como función de enlace, esta función se ajusta bien a la necesidad de aproximar las probabilidades de un resultado binario como función de la variable explicativa, la transformación logística de y a z se obtiene aplicando la función de enlace logit al valor esperado de y , de esta forma esta última queda transformada a una variable continua (z) que a su vez puede usarse como variable dependiente de una regresión que tiene a x como variable explicativa, la diferencia de esta regresión final respecto a una tradicional es que el error no sigue una distribución normal sino logística.

Tipos de Modelos de elección discreta.

Dentro de los modelos de elección discreta existen varios que se usan dependiendo del tipo de variables tanto dependientes como independientes que se tengan:

- Regresión Logística: Se usa cuando al menos una de las variables explicativas es numérica y continua.

- Modelo Logit: Es utilizado cuando todas las variables explicativas son categóricas, en este modelo el coeficiente β está matemáticamente relacionado con el odds ratio o riesgo relativo.
- Modelo Probit: este modelo también se aplica a variables dependientes binarias pero con diferentes supuestos sobre la función de enlace llamada probit y que corresponde a una distribución normal inversa que garantiza que el modelo final estimado sea normal también. La elección entre un modelo logit o probit depende de la distribución de la variable dependiente, si esta se distribuye normal se debe usar el probit en caso contrario puede ser más acertado usar el logit
- Generalizaciones de los Modelos Logit y Probit:
 - Modelos Logit/Probit Ordenados: La variable dependiente no es binaria sino categórica ordinal
 - Modelos Logit/Probit Multinomial: la variable dependiente es categórica pero no ordinal.
 - Modelos Logit/Probit multivariados: Varios modelos de elección discreta son estimados simultáneamente, se usa cuando hay varias variables dependientes.

Aplicación Del Modelo Logit

Razón de aplicación y estructura de datos

Teniendo en cuenta la estructura de los datos recogidos en el experimento piloto en donde se tienen variables cualitativas dicotómicas al lado derecho de la ecuación, es decir como variables explicativas y de igual forma al lado izquierdo, se decide utilizar la técnica de regresión logística en un modelo logit para cada una de las variables independientes dicotomizadas en las 5 piezas musicales creadas, con el fin de observar la existencia de un efecto sobre la satisfacción.

Esta técnica como se mencionó anteriormente, utiliza la regresión logística como medio para explicar el efecto de las variables independientes o explicativas sobre la dependiente, mediante de una función de enlace o transformación logística.

Estadísticos importantes en el Modelo Logit

- Odds ratio (OR) y Riesgo Relativo (RR):

Además de lo anterior, el análisis por el modelo Logit proporciona estadísticos representativos como lo son los odds ratio (OR), utilizados como una medida de riesgo representada mediante el cociente de los odds (ventajas) de cada variación de una variable explicativa determinada.

Estas ventajas (odds) en un ejemplo del caso de estudio, serían en un primer lugar el número de veces que se presenta satisfacción en una pieza determinada frente a las veces que no se presenta para el nivel 1 (modo mayor) de la variable modo musical y en un segundo lugar, corresponderían al número de veces que se presenta satisfacción frente a las veces que no se presenta para el nivel 0 (modo menor) de la variable modo musical. De esta forma se puede observar con mayor facilidad que nivel de la variable explicativa tiene una mayor probabilidad relativa de presentar el suceso determinado que en este caso sería la satisfacción.

De otro lado están también los indicadores de riesgo relativo (RR) calculados de forma distinta y que explican también la probabilidad relativa de que ocurra un suceso determinado en las variaciones binarias de cada variable explicativa. En este caso un ejemplo sería la probabilidad relativa de que haya satisfacción en una pieza musical determinada para valores (1, modo mayor) de la variable modo musical a valores (0, modo musical menor) de dicha variable, lo mismo sería en el caso de la otra variable explicativa, tempo musical, en donde el RR explicaría que probabilidad relativa hay de que haya satisfacción en el caso de tempo rápido frente al tempo lento.

Como se puede ver por los dos caminos de medición se llega al mismo resultado, sin embargo es más usado el odds ratio por la versatilidad que representa en poder ser calculado en tablas 2 x 2 entre otras facilidades (Molinero 2001).

- Estadístico de Wald:

Es un estadístico importante en esta técnica teniendo en cuenta que es el que proporciona información sobre la importancia que tiene cada variable independiente sobre las variaciones de la endógena, es necesario tener en cuenta que escala tiene la variable explicativa para la que se calculará el estadístico ya que dependiendo de si es categórica o no el cálculo del estadístico puede variar.

El uso del estadístico se hace por medio de una prueba de hipótesis que en pocas palabras contrasta la hipótesis nula H_0 de que la variable X_i no es importante para establecer variaciones sobre la variables explicada en la regresión frente a la alternativa H_1 de que la variable X_i si es importante para establecer variaciones en la misma.

- R cuadrado de Nagelkerke y de Cox y Snell:

Estos dos estadísticos representan que parte de la varianza de la variable dependiente es explicada por la regresión que se estime, en el caso de la regresión logística los dos estadísticos son válidos para medir este aspecto, se puede decir que la parte de la varianza explicada de la variable dependiente en el la regresion occila entre estos dos estadísticos, no hay un nivel definido para decir que la varianza explicada es alta o baja esto depende del criterio del investigador.

Aplicación de las regresiones.

Como se mencionaba anteriormente se construye un modelo logit dicotómico para cada una de las variables dependientes y para cada pieza musical, la variable de satisfacción que inicialmente se encuentra medida en una escala del 0 al 9 es recategorizada en una medida de satisfacción binaria para la aplicación del modelo, de igual forma dicha recategorización se

aplica a cada una de las otras variables dependientes para que queden en términos de presencia de satisfacción (1) o ausencia de la misma (0) y así poder generar los correspondientes modelos en cada una de las 5 piezas musicales.

Tenemos entonces las siguientes variables dependientes recategorizadas y expresadas en términos de satisfacción o no satisfacción así:

- Satisfacción:
 - Escala original de 0 a 4 → Insatisfecho con valor 0
 - Escala de 5 al 9 → Satisfecho con valor 1
- Número de Parpadeos:

Para llevar a valores homogéneos y finalmente binarios a esta variable inicialmente se genera el número de parpadeos por minuto en cada pieza, se toma como referencia 4 parpadeos por minuto para definir la variable binaria, esto teniendo en cuenta que en adultos jóvenes las tasas de parpadeos por minuto cuando la persona se encuentra mirando objetos fijos como en este caso con el computador son de entre 3 y 4 parpadeos por minuto (Doughty 2002).

De esta forma la variable binaria queda de la siguiente forma:

- Menos de 4 parpadeos → Insatisfecho con valor 0
- Más de 4 parpadeos → Satisfecho con valor 1
- Latencia o Duración del Parpadeo:

Para llevar esta variable a medidas binarias se tomó como referencia el valor de 300 milisegundos para definir la satisfacción del individuo, teniendo en cuenta que el parpadeo promedio dura aproximadamente 300 milisegundos (Litherland 2009) de este modo se define la variable binaria así:

- Menos de 300 milisegundos → Insatisfecho con valor 0
- Más de 300 milisegundos → Satisfecho con valor 1

- Tiempo de respuesta:

Para llevar esta variable a medidas binarias y teniendo en cuenta que por la naturaleza de la medida no hay parámetros universales ya que es una variable específica de esta prueba, se sacan promedios de la variable para cada pieza musical y dependiendo de si su valor se encontraba sobre el promedio se define la satisfacción así:

- Por debajo del promedio → Insatisfecho con valor 0
- Por encima del promedio → Satisfecho con valor 1

De otro lado, como ya se explicaba anteriormente tenemos las variables Independientes Modo Musical y Tempo Musical, cuyos valores binarios se definen de la siguiente forma:

- Tempo Musical:
 - Rápido → 1
 - Lento → 0
- Modo Musical:
 - Mayor → 1
 - Menor → 0

Una vez definido lo anterior se corre un modelo para cada variable dependiente y en cada una de las piezas musicales y se obtienen los resultados que se muestran a continuación:

Resultados de las regresiones.

Se obtienen los siguientes resultados después de correr las regresiones logísticas de las variables Modo y Tempo versus cada una de las variables dependientes para cada pieza musical.

De forma general se presentan los principales indicadores que en promedio se obtuvieron para cada regresión en las 5 piezas musicales.

Tabla 23 *Estadísticas generales promedio regresiones logísticas.*

Regresiones	R cuadrado de Cox y Snell (Promedio)	R cuadrado de Nagelkerke (Promedio)	% de clasificación Promedio	Pieza Sobresaliente
MyT vs Satisfacción	15,4%	20,6%	67,2	P1
MyT vs Num Parpadeos	20,2%	27,4%	72,9	P3
MyT vs T Respuesta	11,8%	15,9%	70,7	P4
MyT vs Latencia	9,7%	13,1%	71,4	P1

Como se puede observar de forma general no se encuentran porcentajes importantes de varianza explicada en promedio para las 5 piezas musicales, ya que estos se ubican entre el 13% y 28%. A primera vista, la regresión en la que se presentan los mayores porcentajes de varianza explicada de las variables Tempo y Modo Musical es la que tiene como variable dependiente al Número de Parpadeos (variable dicotomizada). En esta regresión se ve que el porcentaje promedio de explicación de la variable dependiente es del 27,4% según el R cuadrado de Nagelkerke, de otro lado se evidencia también que dicha regresión presenta el mayor porcentaje de clasificación correcta con un valor del 72,9 % de casos bien clasificados o predichos, si se analizan por aparte los casos 1 (Hartos Parpadeos) y 0 (pocos parpadeos), el porcentaje promedio de clasificación correcta para los casos que marcan pocos parpadeos es del 81,5% mientras que para los que marcan hartos es del 58,6%, concentrándose así el error más grande en los casos objetivo del modelo, lo que significa una capacidad más reducida del modelo para predecir valores 1 (Hartos Parpadeos) relacionados con la satisfacción.

De otro lado es importante también resaltar que la regresión realizada con la variable satisfacción no presenta los mejores o más contundentes resultados estadísticos para apoyar con evidencia el efecto que en general se espera que tengan las variables independientes analizadas sobre dicha variable independiente. Sin embargo de las 4 regresiones se encuentra en segundo lugar según los valores de sus estadísticos principales ya que en comparación a las

regresiones corridas para las variables Tiempo de Respuesta y Latencia presenta un mayor porcentaje de varianza explicada pero con un error de clasificación mayor.

Por último, se puede ver que de todas las 5 piezas, la número 1 sobresale en 2 de las regresiones por tener los mejores indicadores estadísticos (R cuadrado de Nagelkerke y de Cox y Snell, % de clasificación total y significancia del estadístico de Wald para las 2 variables independientes Tempo y Modo Musical).

Profundizando en el análisis para cada una de las regresiones tenemos lo siguiente:

- Modo y Tempo VS Satisfacción:

Tabla 24 *R cuadrado modo y tempo vs satisfacción*

Pieza Musical	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke	Prueba Hosmer y Lemeshow		
				Chi cuadrado	gl	Sig.
1	225,543	,227	,303	21,774	2,0	,000
2	239,515	,167	,223	19,941	2,0	,000
3	245,821	,145	,194	22,992	2,0	,000
4	250,976	,120	,160	18,842	2,0	,000
5	240,401	,109	,149	25,476	2,0	,000

Como se puede observar la pieza 1 es la que presenta los mayores niveles en cuanto a la varianza explicada por las variables Modo y Tempo Musical respecto a la de Satisfacción según el estadístico de Nagelkerke, sin embargo es de apenas el 30% por lo tanto en todas las piezas musicales se observan valores más bajos que este. También se puede ver que para todas las regresiones con la variable de satisfacción como dependiente se tienen modelos con porcentajes de explicación de varianza significativos si se observan los valores de significancia de las pruebas de Hosmer y Lemeshow que se encuentran en todos los casos por debajo de 0,05.

Tabla 25 matriz de clasificación modo y tempo vs satisfacción

Pieza Musical	Observado		Pronosticado		
			Satisfacción Dicotómica		Porcentaje correcto
			Insatisfecho	Satisfecho	
1	Satisfacción Dicotómica con la Pieza	Insatisfecho	70	27	72,2
		Satisfecho	30	73	70,9
	Porcentaje global				71,5
2	Satisfacción Dicotómica con la Pieza	Insatisfecho	68	38	64,2
		Satisfecho	32	62	66,0
	Porcentaje global				65,0
3	Satisfacción Dicotómica con la Pieza	Insatisfecho	66	35	65,3
		Satisfecho	34	65	65,7
	Porcentaje global				65,5
4	Satisfacción Dicotómica con la Pieza	Insatisfecho	113	13	89,7
		Satisfecho	37	37	50,0
	Porcentaje global				75,0
5	Satisfacción Dicotómica con la Pieza	Insatisfecho	30	62	32,6
		Satisfecho	20	88	81,5
	Porcentaje global				59,0

Se puede ver en general porcentajes de clasificación correctos mayores al 65% en promedio, sobresalen los valores de las piezas 1 y 4 siendo esta última la de mayor porcentaje con el 75%, sin embargo, encontramos en un segundo lugar a la pieza 1 con resultados de clasificación correctos más homogéneos para casos de satisfacción y no satisfacción ya que en la pieza 4 los casos correctos se concentran más en los satisfechos que en los insatisfechos.

Tabla 26 Significancia variables modo y tempo vs satisfacción

Pieza	Variables	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
1	Modo(1)	-1,994	,336	35,245	1	,000	,136	,071	,263
	Tempo(1)	-1,182	,336	12,389	1	,000	,307	,159	,592
	Constante	1,666	,316	27,759	1	,000	5,293		
2	Modo(1)	-,808	,305	7,024	1	,008	,446	,245	,810
	Tempo(1)	-1,293	,305	17,953	1	,000	,275	,151	,499
	Constante	,913	,269	11,561	1	,001	2,493		
3	Modo(1)	-1,369	,311	19,368	1	,000	,254	,138	,468
	Tempo(1)	-1,048	,311	11,356	1	,001	,351	,191	,645
	Constante	1,185	,282	17,688	1	,000	3,271		
4	Modo(1)	-,762	,314	5,896	1	,015	,467	,252	,863
	Tempo(1)	-1,289	,317	16,533	1	,000	,276	,148	,513
	Constante	,422	,258	2,684	1	,101	1,526		
5	Modo(1)	-1,457	,319	20,907	1	,000	,233	,125	,435
	Tempo(1)	-1,215	,318	14,545	1	,000	,297	,159	,554
	Constante	1,529	,301	25,745	1	,000	4,615		

Como se puede observar en términos generales para todas las piezas musicales se tienen valores de significancia menores a 0,05 sugiriendo así que las variables independientes si tienen poder explicativo sobre la dependiente en todas las regresiones que tienen la variable Satisfacción como dependiente, también se observan valores de $\exp(B)$ lejanos de 1 en concordancia con lo anterior, especialmente en la pieza 1 en donde el Modo es la más explicativa. Sin embargo son valores menores que 1 que sugieren evidencia estadística de una relación inversa entre las variables independientes y la dependiente lo que estaría un poco en discordancia con los resultados de las matrices de clasificación.

- Modo y Tempo VS Número de Parpadeos:

Tabla 27 *R cuadrado modo y tempo vs número de parpadeos*

Pieza Musical	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke	Prueba Hosmer y Lemeshow		
				Chi cuadrado	gl	Sig.
1	186,131	,337	,456	4,747	2,0	,093
2	214,479	,194	,268	11,298	2,0	,004
3	223,494	,186	,253	8,198	2,0	,017
4	235,188	,174	,233	1,521	2,0	,467
5	250,761	,118	,158	,144	2,0	,930

La bondad de ajuste para estas regresiones no es tan buena como la de las anteriores ya que solo para las piezas 2 y 3 se observa un porcentaje de explicación de varianza de la variable dependiente, que en este caso es el número de parpadeos, significativo, de acuerdo con los estadísticos de Cox y Snell y Nagelkerke la regresión de la pieza 2 constituye un modelo que explica entre un 19,4 y un 26,8 de la varianza del Número de Parpadeos y aunque es el segundo valor más grande si es significativo, contrario a lo que sucede con la pieza 1 que tiene un porcentaje de explicación mayor (entre el 33,7 y el 45,6) pero no significativo.

Tabla 28 R matriz de clasificación modo y tempo vs número de parpadeos.

Pieza Musical	Observado		Pronosticado		
			Número de Parpadeos Dicotómico		Porcentaje correcto
			Pocos	Hartos	
1	Número de Parpadeos Dicotómico	Pocos	63	29	68,5
		Hartos	37	71	65,7
	Porcentaje global				67,0
2	Número de Parpadeos Dicotómico	Pocos	76	38	66,7
		Hartos	24	62	72,1
	Porcentaje global				69,0
3	Número de Parpadeos Dicotómico	Pocos	110	11	90,9
		Hartos	40	39	49,4
	Porcentaje global				74,5
4	Número de Parpadeos Dicotómico	Pocos	113	12	90,4
		Hartos	37	38	50,7
	Porcentaje global				75,5
5	Número de Parpadeos Dicotómico	Pocos	119	12	90,8
		Hartos	31	38	55,1
	Porcentaje global				78,5

Respecto a las matrices de clasificación es la pieza 5 la que cuenta con el mayor porcentaje global de clasificación correcto con un 78,5% en la cual también se observa un poder de predicción mayor para los valores (0) que para los valores (1) de la variable objetivo con un 90,8% y un 55,1% para “Pocos” y “Hartos” respectivamente. En un segundo y tercer lugar se encuentran las piezas 4 y 3 respectivamente con porcentajes globales de clasificación correcta de 75,5% y 74,5% respectivamente.

Tabla 29 Significancia variables modo y tempo vs número de parpadeos

Pieza	Variables	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
1	Modo(1)	-,366	,304	1,452	1	,228	,694	,383	1,258
	Tempo(1)	-1,439	,304	22,387	1	,000	,237	,131	,430
	Constante	1,085	,275	15,576	1	,000	2,960		
2	Modo(1)	-,882	,320	7,567	1	,006	,414	,221	,776
	Tempo(1)	-1,716	,323	28,311	1	,000	,180	,095	,338
	Constante	,954	,276	11,981	1	,001	2,596		
3	Modo(1)	-1,624	,397	16,754	1	,000	,197	,091	,429
	Tempo(1)	-2,908	,415	49,110	1	,000	,055	,024	,123
	Constante	1,584	,340	21,756	1	,000	4,877		
4	Modo(1)	-,887	,331	7,179	1	,007	,412	,215	,788
	Tempo(1)	-1,875	,340	30,383	1	,000	,153	,079	,299
	Constante	,739	,272	7,406	1	,007	2,095		
5	Modo(1)	-1,139	,343	11,044	1	,001	,320	,164	,627
	Tempo(1)	-1,897	,355	28,597	1	,000	,150	,075	,301
	Constante	,700	,274	6,528	1	,011	2,013		

Se observan niveles de significancia por debajo del 0,05 en general para todos los casos excepto para la variable Modo en la regresión de la pieza 1, al igual que en la anterior (Satisfacción como dependiente) en este caso los valores de exp(B) también se encuentran todos debajo de 1 implicando relaciones inversas entre las variables dependientes y la independiente contrario nuevamente a la relación esperada.

- Modo y Tempo VS Tiempo de Respuesta:

Tabla 30 R cuadrado Modo y Tempo VS Tiempo de Respuesta

Pieza Musical	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke	Prueba Hosmer y Lemeshow		
				Chi cuadrado	gl	Sig.
1	236,635	,168	,225	16,564	2,0	,000
2	240,401	,147	,197	24,794	2,0	,000
3	247,091	,108	,146	10,695	2,0	,005
4	254,870	,097	,130	12,140	2,0	,002
5	259,294	,073	,097	6,442	2,0	,040

Respecto a estas Regresiones el modelo con mejor porcentaje de varianza explicada de la variable explicada Tiempo de respuesta corresponde al de la pieza musical 1 con un

porcentaje de entre el 17% y el 22,5% de acuerdo a los estadísticos de Cox y Snell y Nagelkerke respectivamente, también se puede apreciar que los R cuadrado para todas las piezas musicales de estas regresiones son significativos lo que hace confiables los valores de los R cuadrado contenidos en la Tabla 10.

Tabla 31 *R* matriz de clasificación modo y tempo VS tiempo de respuesta.

Pieza Musical	Observado		Pronosticado		
			Tiempo de Respuesta Dicotómico		Porcentaje correcto
			Lento	Rápido	
1	Tiempo de Respuesta Dicotómico	Lento	106	13	89,1
		Rápido	44	37	45,7
	Porcentaje global				71,5
2	Tiempo de Respuesta Dicotómico	Lento	69	41	62,7
		Rápido	31	59	65,6
	Porcentaje global				64,0
3	Tiempo de Respuesta Dicotómico	Lento	109	7	94,0
		Rápido	41	43	51,2
	Porcentaje global				76,0
4	Tiempo de Respuesta Dicotómico	Lento	107	7	93,9
		Rápido	43	43	50,0
	Porcentaje global				75,0
5	Tiempo de Respuesta Dicotómico	Lento	98	14	87,5
		Rápido	52	36	40,9
	Porcentaje global				67,0

Como se muestra en la tabla, en general el poder de clasificación de las diferentes regresiones para cada pieza musical es bueno a nivel global ya que se encuentran todos por encima del 60%, las tres mejores piezas respecto a estos indicadores son la 3,4 y 1 con porcentajes arriba del 70% siendo la 3 la del mayor con un 76%, sin embargo, los resultados de clasificación parecen ser contrarios a lo esperado teniendo en cuenta que en todas las piezas el porcentaje correcto de clasificación de las variables independientes se concentra más para el resultado 0, “Lento” de la variable Tiempo de Respuesta, mientras que para el valor “Rápido” el porcentaje de clasificación correcto es más bajo en todos los casos.

Tabla 32 Significancia variables modo y tempo VS tiempo de respuesta.

Pieza Musical	Variables	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
1	Modo(1)	-,691	,308	5,048	1	,025	,501	,274	,916
	Tempo(1)	-1,281	,309	17,152	1	,000	,278	,151	,509
	Constante	,552	,259	4,554	1	,033	1,737		
2	Modo(1)	-1,191	,301	15,641	1	,000	,304	,168	,548
	Tempo(1)	-,620	,301	4,257	1	,039	,538	,298	,969
	Constante	,683	,260	6,904	1	,009	1,980		
3	Modo(1)	-1,206	,317	14,497	1	,000	,299	,161	,557
	Tempo(1)	-1,289	,317	16,533	1	,000	,276	,148	,513
	Constante	,867	,271	10,234	1	,001	2,379		
4	Modo(1)	-1,400	,322	18,904	1	,000	,247	,131	,464
	Tempo(1)	-1,318	,322	16,768	1	,000	,268	,143	,503
	Constante	1,017	,278	13,350	1	,000	2,766		
5	Modo(1)	-,772	,297	6,754	1	,009	,462	,258	,827
	Tempo(1)	-,854	,297	8,248	1	,004	,426	,238	,763
	Constante	,553	,255	4,678	1	,031	1,738		

Nuevamente se pueden observar valores de significancia por debajo de 0,05 en todos los casos para cada una de las regresiones, al igual que en los anteriores análisis, Exp (B) es menor que 1 indicando una relación inversa entre la variable dependiente y las explicativas es decir que el tiempo de respuesta tiende a ser rápido cuando el modo y el tempo tienen valores 0 (menor y lento respectivamente).

- Modo y Tempo VS Latencia:

Tabla 33 R cuadrado modo y tempo vs latencia

Pieza Musical	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke	Prueba Hosmer y Lemeshow		
				Chi cuadrado	gl	Sig.
1	238,813	,121	,165	7,952	2,0	,019
2	243,927	,119	,161	7,415	2,0	,025
3	246,194	,093	,126	13,487	2,0	,001
4	250,154	,091	,123	6,036	2,0	,049
5	259,126	,060	,080	10,678	2,0	,005

Los estadísticos muestran que la pieza 1 es aquella que cuenta con el mayor nivel de varianza explicada con valores entre el 12,1% y el 16,5%, también se ve en la tabla que según la prueba de Hosmer y Lemeshow los R cuadrado que se ven en cada una de las regresión es

para la variable explicada Latencia, son significativos, podemos ver en último lugar a la pieza 5 que apenas cuanta con valores de varianza explicada de entre el 6% y el 8%.

Tabla 34 *Matriz de clasificación modo y tempo vs latencia*

Pieza Musical	Observado		Pronosticado		
			Duración del Parpadeo Dicotómica		Porcentaje correcto
			Poca	Harto	
1	Duración del Parpadeo Dicotómica	Poca	111	14	88,8
		Harto	39	36	48,0
	Porcentaje global				73,5
2	Duración del Parpadeo Dicotómica	Poca	105	15	87,5
		Harto	45	35	43,8
	Porcentaje global				70,0
3	Duración del Parpadeo Dicotómica	Poca	110	14	88,7
		Harto	40	36	47,4
	Porcentaje global				73,0
4	Duración del Parpadeo Dicotómica	Poca	107	13	89,2
		Harto	43	37	46,3
	Porcentaje global				72,0
5	Duración del Parpadeo Dicotómica	Poca	102	15	87,2
		Harto	48	35	42,2
	Porcentaje global				68,5

Se puede ver en la matriz de clasificación que los porcentajes correctos de clasificación global están alrededor del 70%, se destacan los de las piezas 1 y 3 con 73,5% y 73% respectivamente, sin embargo en general la clasificación correcta se concentra en los valores 0 de la variable de Latencia es decir en “poca”, mientras que para los valores 1 que representan una duración larga los porcentajes de clasificación correcta son bajos en general.

Tabla 35 Significancia variables modo y tempo vs latencia

Pieza Musical	Variables	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
1	Modo(1)	-1,086	,317	11,698	1	,001	,338	,181	,629
	Tempo(1)	-1,173	,318	13,602	1	,000	,310	,166	,577
	Constante	,543	,261	4,320	1	,038	1,721		
2	Modo(1)	-,980	,306	10,264	1	,001	,375	,206	,684
	Tempo(1)	-,896	,306	8,593	1	,003	,408	,224	,743
	Constante	,490	,257	3,646	1	,056	1,632		
3	Modo(1)	-1,001	,310	10,443	1	,001	,368	,200	,674
	Tempo(1)	-,915	,309	8,748	1	,003	,401	,218	,734
	Constante	,415	,256	2,629	1	,105	1,515		
4	Modo(1)	-1,182	,313	14,293	1	,000	,307	,166	,566
	Tempo(1)	-1,013	,312	10,548	1	,001	,363	,197	,669
	Constante	,635	,262	5,866	1	,015	1,886		
5	Modo(1)	-,648	,297	4,763	1	,029	,523	,292	,936
	Tempo(1)	-,814	,297	7,491	1	,006	,443	,247	,794
	Constante	,365	,252	2,103	1	,147	1,441		

Al igual que en los casos anteriores se observan valores de Exp (B) menores que 1 para todos los casos sugiriendo nuevamente relaciones inversas entre las variables independientes y la explicada, es decir, que Modos Mayores y Tempos Rápidos generan valores de latencia bajos, contrario a lo esperado.

Conclusiones de las regresiones.

En general se puede ver que para todas las regresiones los estadísticos de Cox y Snell y Nagelkerke son mayores al 9,7% y 13,1% en promedio respectivamente y llegan hasta un 20% y 27%, como en el caso de los valores para las regresiones Modo y Tempo versus Número de parpadeos, se observa igualmente que son estas regresiones las que también cuentan con los porcentajes de clasificación promedio más altos. En todos los casos dichos estadísticos son siempre significativos lo que les da confiabilidad.

De otro lado, las regresiones del Modo y Tempo Vs la variable Satisfacción aunque no estuvieron entre los indicadores promedio más bajos respecto a sus R cuadrado y porcentajes de clasificación, tampoco fueron los mejores como podría esperarse.

De forma general se destacó la pieza 1, que tuvo los mejores indicadores estadísticos en dos de los grupos de regresiones respecto a R Cuadrados y porcentajes de clasificación

correctos: Modo y Tempo vs Satisfacción y Modo y Tempo vs Latencia mientras que piezas como la 5 y 3 estuvieron en todas las regresiones en los últimos lugares especialmente la pieza la 5.

Finalmente se observa que ,aunque las regresiones muestran en todos sus casos excepto en las constantes, que las variables Tempo y Modo de forma general son significativas, los ODDS Ratio o Exp (B) de cada variable son menores a 1, indicando que no hay una ventaja en la satisfacción cuando los valores de Modo y Tempo son 1, es decir que la probabilidad de que se presente satisfacción es mayor cuando los valores de Modo y Tempo son 0 que en caso contrario (1) opuesto a lo que se esperaría de acuerdo a la teoría Scherer y Zetner (2001).

Árboles De Decisión

Los Árboles de decisión son metodologías estadísticas utilizadas para la predicción y clasificación en el ámbito de la inteligencia artificial. A partir de un conjunto de datos se construyen diagramas que representan y categorizan una serie de eventos que ocurren de forma sucesiva para la resolución de un problema.

Estos árboles tienen entradas que representan los atributos de una determinada situación o problema a resolver a partir de las cuales, por medio de un conjunto de reglas lógicas, generan una respuesta del problema como consecuencia de dichas entradas, tanto estas últimas como las salidas, pueden tomar valores discretos o continuos, en cuyos casos se habla respectivamente de algoritmos de clasificación o regresión dependiendo del caso.

En términos generales los árboles de decisión tienen varios tipos de nodos dentro de los que se encuentran los de probabilidad, internos y hojas. Los primeros indican las probabilidades de que ocurra un evento aleatorio según la naturaleza del problema, los internos contienen test sobre los valores que toma alguna de las propiedades evaluadas, por su parte los nodos hojas son los resultados que devuelve el árbol cuyas ramas finalmente representan los diferentes caminos u opciones que se tienen dependiendo la decisión que se tome.

Estas son técnicas muy usadas hoy en día ya que son metodologías muy flexibles que permiten manejar un gran número de variables con complejas interacciones entre ellas y cuyos resultados terminan siendo de fácil interpretación y entendimiento para cualquier persona (Pérez 2013). Se encuentran dentro de las utilidades de análisis para las que se usan dichas técnicas las siguientes:

- Estratificación: Asigna casos a una categoría entre varias.
- Segmentación: Identifica individuos que pueden pertenecer a un grupo específico según sus características.
- Predicción: Crea reglas que utiliza para predecir eventos futuros como la probabilidad de que una persona caiga en mora un crédito o el valor en que se venderá algún bien o artículo.
- Reducción de datos y clasificación de variables: A partir de un gran número de variables, selecciona un subconjunto útil de predictores que para la creación de un modelo paramétrico formal.
- Identificación de interacciones: Identifica relaciones que pertenecen a subgrupos específicos y las plasma en un modelo formal.
- Fusión de categorías y discretización de variables continuas: Recodifica las categorías de grupo de los predictores y las variables continuas con pérdida mínima de información.

Los árboles de decisión son técnicas que tienen aplicaciones en diferentes campos dentro de los cuales se pueden encontrar las siguientes:

- Envío de Correos: Se pueden usar los árboles para identificar grupos de usuarios con alta tasa de respuesta.

- Modelos Scoring de Crédito: Toma de decisiones más acertadas de colocación de créditos al usar el historial crediticio para identificar grupos con mayor probabilidad de pago.
- Recursos Humanos: Afinar los procesos actuales mediante el entendimiento de las reglas pasadas de contratación.
- Análisis de mercado: Identificación de variables representativas que segmenten los consumidores: geográficas, precios, características, etc.
- Control de Calidad: determinar reglas que permitan identificar de forma general cuando un producto puede ser defectuoso.
- Estudio de políticas: Generación de reglas de decisión en políticas de contratación.
- Salud: Identificación de variables que contribuyan a lograr mejores resultados en este campo.

Como se menciona anteriormente los árboles son técnicas muy versátiles que se pueden aplicar en muchos campos, finalmente se mencionan otras de las ventajas o desventajas, dependiendo del punto de donde se mire, que tienen estas técnicas.

- Son técnicas diseñadas para trabajar con grandes números de variables como sucede especialmente con el algoritmo CHAID.
- Capturan relaciones que los modelos lineales no son capaces de capturar, específicamente relaciones no lineales e interacciones complejas).
- Estas técnicas no generan ecuaciones generales como sucede con las regresiones, sin embargo, si crea reglas de decisión.

Métodos de crecimiento de los árboles de decisión.

Dependiendo de lo que se requiera como los tipos de datos, tipos de medida de las variables, si se quiere usar o no criterios estadísticos al momento de escoger a partir de las variables independientes, si se quiere o no la incorporación de costos de clasificación errónea

y probabilidades a priori o dependiendo también de si se desea hacer más rápido o eficiente el tiempo de procesamiento de la máquina, existen 4 métodos dentro de los cuales escoger para buscar un mejor ajuste a los datos y a las necesidades del investigador estos corresponden a los siguientes:

- CHAID (Chi-Square Automatic Interaction detector): Este método consiste en un algoritmo eficaz y rápido de exploración de datos que crea perfiles y segmentos según el resultado esperado, este algoritmo permite detectar automáticamente interacciones mediante la distribución Chi-cuadrado. En cada paso el algoritmo CHAID va eligiendo la variable independiente que representa la interacción más fuerte con la variable dependiente, en esta técnica las categorías de cada predictor se funden si no son significativamente diferentes respecto a la variable explicada.
- CHAID exhaustivo: Corresponde con a una modificación del algoritmo CHAID que examina todas las divisiones posibles de cada variable explicativa o predictora y las trata a todas por igual independientemente del tipo y número de categorías.
- Árboles de Clasificación y Regresión (CRT-Clasification and regression trees): el CRT es un algoritmo binario que hace particiones de los datos generando subconjuntos lo más homogéneos posibles respecto a la variable dependiente.
- QUEST: (Quick, unbiased, efficient, statistical, tree): Es un algoritmo estadístico que selecciona variables eficientemente a la vez que evita el sesgo que presentan otros métodos que favorecen predictores con muchas categorías (Berlanga et al., 2013), este método solo puede usarse cuando la variable dependiente es nominal.

Tabla 36 Comparación de métodos para árboles de decisión

Característica	CHAID	CHAID Exhaustivo	C&RT	QUEST
Tipo de Partición	Múltiple	Múltiple	Binaria	Binaria
Dependiente Continua	Sí	Sí	Sí	No
Predictoras Continuas	Sí (*)	Sí (*)	Sí	Sí
Coste de Mala Clasificación (Crecimiento del Arbol)	No	No	Sí	Sí
Pruebas Estadísticas (Selección del Predictor)	Sí	Sí	No	Sí
Pruebas Estadísticas (Particionar)	Sí	Sí	No	No
Velocidad	Moderada	Moderada	Lento	Moderada/Lento
Utiliza A priori?	No	No	Sí	Sí
Valores Faltantes para los Predictores Usados?	Si, como una categoría	Si, como una categoría	No, Sustitutos usados para partición	No, Sustitutos usados para partición

Fuente: Jorge Del Río L, Introducción a SPSS Árboles de Clasificación

Como se puede observar en la tabla comparativa, depende de lo que requiera el investigador un método u otro puede ajustarse mejor a sus expectativas, en este caso teniendo en cuenta que los datos que se están trabajando son de corte cualitativo tanto en las variables dependientes como en las independientes, que no es un número grande de registros ni de variables, que no se está estableciendo ninguna probabilidad a priori y que se busca la utilización de pruebas estadísticas para particionar con el objetivo de obtener interrelaciones estadísticamente significativas y más robustas, se ha elegido la metodología CHAID como la más adecuada a los propósitos de este trabajo, por lo que en seguida se profundizará un poco más en los beneficios, estadísticos y medidas que utiliza dicha metodología y su aplicación en este caso.

Árboles de Decisión CHAID.

La técnica CHAID Chi-square Automatic Interaction Detector (Detector Automático de Interacción Chi-cuadrado) es un método usado principalmente para identificar y estudiar las relaciones entre una variable dependiente y un gran número de variables independientes, el método selecciona el sistema de predictores y las interacciones que predicen de forma óptima

la medida de la variable dependiente, en Segmentación de mercados es usado para obtener tipologías y perfiles de los consumidores de un producto o usuarios de un servicio.

Históricamente esta técnica tiene sus fundamentos en las técnicas “Classification and Regression Trees” (CART) y “Automatic Interaction Detection”, en general funciona como una regresión múltiple para variables nominales, categóricas, ordinales, discretas o discontinuas como los son por ejemplo el sexo, nivel socioeconómico, ocupación, ciudad, etc. en donde existe al menos una variable independiente que predice a la variable dependiente, se diferencia de una regresión múltiple tradicional en que esta última trabaja generalmente con variables de intervalo o razón como lo son el ingreso, el tiempo, el peso, etc.

Una de las ventajas y aportes más importantes de esta técnica es la eficiencia en términos de tiempo ya que ahorra tiempo al investigador al evitar la necesidad de que tenga que analizar muchas tabulaciones cruzadas para encontrar alguna relación importante entre variables, esto clasifica al árbol CHAID, como una técnica con un gran potencial de análisis para la toma eficiente de decisiones.

Esta técnica divide en primer lugar al total de la población en dos o más grupos de acuerdo a las categorías del mejor predictor de la variable dependiente y posteriormente sigue subdividiendo estos grupos en otros más pequeños según otros predictores que siguen siendo estadísticamente significativos, hasta que no se encuentren más predictores de este tipo o hasta que se cumpla alguna regla definida para detener el crecimiento del árbol, de otro lado fusiona las categorías de una variable predictora cuando no son estadísticamente diferentes, la combinación de estos dos procedimientos asegura que los casos pertenecientes al mismo segmento sean homogéneos respecto al criterio de segmentación y que los que pertenecen a diferentes segmentos sean heterogéneos entre sí.

Como técnica de agrupamiento para crear los grupos en un análisis, utiliza criterios de variables predictoras a diferencia de la técnica de Cluster que utiliza criterios de distancias,

esto permite que los resultados de un agrupamiento por CHAID puedan usarse con otras muestras, contrario a lo que sucede con los que genera un análisis de Cluster.

La técnica CHAID es una modificación propuesta por Kass (1994) , al ya existente algoritmo de clasificación AID, (Automatic Interaction Detection), esta modificación consistió en hacer un algoritmo de clasificación que utilizaba una función secuencial y un procedimiento de división basado en la prueba Ji cuadrado que ahorra mucho tiempo de procesamiento en las máquinas en lugar de buscar relaciones entre variables a través del análisis de múltiples tablas de contingencia y además de esto solucionaba el problema que tenían en general los árboles de no tener alguna regla definida para detener su crecimiento y división para no terminar con nodos con un solo caso ya que con este nuevo procedimiento basado en la prueba Ji cuadrado que utiliza la falta de significancia como argumento para detener la subdivisión de las ramas (Alpizar 1994).

El análisis CHAID tiene dos componentes básicos a saber: las variables predictivas, que se utilizan para definir los segmentos y en donde la flexibilidad del análisis permite usar cualquier tipo de ellas: demográficas, psicográficas, estilo de vida etc. y de otro lado está el criterio o variable dependiente que construye el modelo de segmentación que está restringido a una sola variable que debe ser categórica u ordinal.

General mente la variable dependiente para este análisis es dicotómica sin embargo también puede trabajar con variables dependientes que tengan más de dos categorías (variables politomas) las variables continuas se pueden convertir en categóricas al agruparlas en rangos de acuerdo a lo que se requiera, como por ejemplo la edad: de 18 a 24 o de 24 a 40 etc.

La variable de frecuencia representa el número de veces que las observaciones están en un grupo en particular, esta debe ser igual al total de las observaciones de la muestra.

Esta técnica permite el uso de una variable ponderadora en caso de requerirse, esta variable le da tratamientos desiguales a los casos de un conjunto de casos que pueden ser

ponderados para representar mejor a la población o en el caso de segmentación de mercados representar mejor las particiones reales del mercado, cuando se utiliza este tipo de variable la contribución de un caso en el análisis se pondera proporcionalmente a las unidades de población que el caso represente en la muestra.

Existen tres tipos de variables predictoras en el análisis CHAID, monofónicos, libres o flotantes dependiendo de los que se tienen y su clasificación se verá afectado el algoritmo CHAID respecto a la fusión de categorías y por consiguiente el cálculo de los niveles de significancia, en el caso de predictores dicotómicos (Con dos categorías), como es el caso del análisis realizado en este trabajo, no es necesaria esta clasificación ya que el algoritmo CHAID trata a este tipo de predictores de la misma forma.

Cuando una variable es monotónica se pueden combinar sus categorías solo si estas son adyacentes entre sí como sucede por ejemplo con la edad, respecto a las variables libres, CHAID puede fusionar categorías incluso si estas no son adyacentes, todas las variables nominales cuyas categorías no tienen un orden natural deben ser definidas como libres.

El algoritmo que utiliza esta técnica tiene 3 pasos o etapas principales a saber: fusión, división y paro.

1. **Etapas de fusión:** utiliza cinco pasos para la unión de categorías en cada uno de las predicciones x_1, x_2, \dots, x_K
 - a. Realiza tabulación cruzada de dos vías con la variable dependiente.
 - b. Para cada par de categorías que es factible de fusión, calcula la estadística Ji-cuadrada, con la que prueba la independencia del par de categorías con la variable dependiente.
 - c. En seguida calcula el p valor de la prueba.

- d. Con el nivel de significancia de las pruebas, comprueba, para variables con más de 2 categorías, que si la que es predictora resulta relevante y debe ser separada de las otras, si esto resulta así efectivamente se separa.
 - e. Fusiona las categorías que tengan pocas observaciones.
 - f. Computa el valor ajustado de Bonferroni que es un test de comparaciones múltiples que reduce el error de comparar de dos en dos (Goldman, 2008), para fusionar las categorías.
2. **Etapa de División:** de acuerdo a las predicciones y a sus p valor mira si estos son significativos, si esto sucede divide el grupo por el p valor menos significativo, en caso de que no haya valores p significativos el grupo no se divide.
 3. **Paro:** Vuelve al paso 1 para efectuar los pasos de fusión sobre el subgrupo creado, antes de dividirlo, finalmente se detiene cuando haya analizado todos los grupos.

Resultados árboles CHAID.

En el presente análisis, por las razones anteriormente explicadas, se decide la utilización de la técnica CHAID, el método de abordar este análisis se alinea con el capítulo anterior de las regresiones logísticas, de este modo se ilustrarán los resultados de los árboles CHAID corridos para cada pieza musical y también para cada una de las variables independientes a saber: Satisfacción Dicotómica, Número de parpadeos dicotómico, Duración del Pardeo Dicotómico Y tiempo de Respuesta Dicotómico.

Teniendo en cuenta que estamos ante un algoritmo CHAID de tipo dicotómico, cada una de las ramas de los árboles de decisión que se corren contiene dos nodos ya que las variables independientes Modo y Tempo musical tienen 2 categorías cada una (Mayor y Menor en la primera y Rápido y Lento en la segunda), de esta forma se mostrarán tablas que agrupan para cada una de las variables dependientes que se analizan (Satisfacción, Número de parpadeos, Tiempo de Respuesta y Latencia dicotómicas) los 5 árboles de decisión.

Teniendo en cuenta lo anterior, en cada pieza musical la tabla del árbol de decisión correspondiente lo representará con sus nodos así: el nodo 0 es el inicio con el total de los casos de la muestra de entrenamiento y la representación de los porcentajes de los casos que corresponden a la categoría pronosticada y los que no, de aquí en adelante los diferentes niveles del árbol estarán conformados por dos nodos, cuyos números corresponden siempre al nodo de la izquierda para el caso del nodo con el menor número y a la de la derecha para el del mayor, así por ejemplo el segundo nivel de un árbol estará conformado por los nodos 1 y 2 de los cuales el 1 es el de la izquierda y 2 el de la derecha, posteriormente el tercer nivel del árbol estará conformado por los nodos 3 y 4 en donde el 3 es el de la izquierda y el 4 el de la derecha, cada uno se dividirá según la categoría que mejor represente la variable predicha.

- Modo y Tempo VS Satisfacción:

Tabla 37 Árboles Chaid modo y tempo vs satisfacción

Arbol	Nodo	Insatisfech		Satisfecho		Total		Categoría pronosticada	Nodo parental	Variable independiente primaria				
		N	%	N	%	N	%			Variable	Sig.a	Chi-cuadrado	gl	Valores de segmentación
Arbol Pieza 1	0	62	46%	74	54%	136	100%	Satisfecho						
	1	16	24%	51	76%	67	49%	Satisfecho	0	Modo Musical Dicotómico	0,00	26	1	Modo Mayor
	2	46	67%	23	33%	69	51%	Insatisfecho	0	Modo Musical Dicotómico	0,00	26	1	Modo Menor
	3	15	42%	21	58%	36	26%	Satisfecho	1	Tempo Musical Dicotómica	0,00	16	1	Lento
Arbol Pieza 2	4	1	3%	30	97%	31	23%	Satisfecho	1	Tempo Musical Dicotómica	0,00	16	1	Rápido
	0	81	54%	69	46%	150	100%	Insatisfecho						
	1	33	44%	42	56%	75	50%	Satisfecho	0	Modo Musical Dicotómico	0,01	6	1	Modo Mayor
	2	48	64%	27	36%	75	50%	Insatisfecho	0	Modo Musical Dicotómico	0,01	6	1	Modo Menor
Arbol Pieza 3	3	27	71%	11	29%	38	25%	Insatisfecho	1	Tempo Musical Dicotómica	0,00	24	1	Lento
	4	6	16%	31	84%	37	25%	Satisfecho	1	Tempo Musical Dicotómica	0,00	24	1	Rápido
	0	66	49%	70	51%	136	100%	Satisfecho						
	1	40	62%	25	38%	65	48%	Insatisfecho	0	Modo Musical Dicotómico	0,00	9	1	Modo Menor
Arbol Pieza 4	2	26	37%	45	63%	71	52%	Satisfecho	0	Modo Musical Dicotómico	0,00	9	1	Modo Mayor
	3	3	8%	33	92%	36	26%	Satisfecho	2	Tempo Musical Dicotómica	0,00	28	1	Rápido
	4	23	66%	12	34%	35	26%	Insatisfecho	2	Tempo Musical Dicotómica	0,00	28	1	Lento
	0	94	67%	47	33%	141	100%	Insatisfecho						
Arbol Pieza 5	1	41	59%	29	41%	70	50%	Insatisfecho	0	Modo Musical Dicotómico	0,04	4	1	Modo Mayor
	2	53	75%	18	25%	71	50%	Insatisfecho	0	Modo Musical Dicotómico	0,04	4	1	Modo Menor
	3	31	84%	6	16%	37	26%	Insatisfecho	1	Tempo Musical Dicotómica	0,00	22	1	Lento
	4	10	30%	23	70%	33	23%	Satisfecho	1	Tempo Musical Dicotómica	0,00	22	1	Rápido
Arbol Pieza 5	0	65	45%	81	55%	146	100%	Satisfecho						
	1	18	26%	51	74%	69	47%	Satisfecho	0	Modo Musical Dicotómico	0,00	18	1	Modo Mayor
	2	47	61%	30	39%	77	53%	Insatisfecho	0	Modo Musical Dicotómico	0,00	18	1	Modo Menor
	3	17	59%	12	41%	29	20%	Insatisfecho	1	Tempo Musical Dicotómica	0,00	31	1	Lento
Arbol Pieza 5	4	1	3%	39	98%	40	27%	Satisfecho	1	Tempo Musical Dicotómica	0,00	31	1	Rápido

Al correr los árboles CHAID para la variable dependiente Satisfacción Dicotómica se puede observar que en 3 de las piezas musicales (1,3 y 5), la mayoría de los casos se encuentran clasificados dentro de la categoría Satisfecho (1), sin embargo, en todas las piezas musicales las variables independientes se comportan de acuerdo a lo esperado en sus categorías ya que en todos los casos el porcentaje de observaciones para las categorías Mayor (1) de la variable Modo y Rápido (1) de la variable Tempo, es mayor que el de los valores 0 en las mismas variables, (Menor para la primera y Lento para la segunda).

De otro lado también se observa que la variable que más impacto tiene sobre la categoría objetivo “Satisfecho” (1) de la variable dependiente es la de Modo Musical, esto teniendo en cuenta que en todos los árboles el segundo nivel está representado por esta variable en donde a su vez la categoría Mayor (1) se convierte en objetivo según la prueba ji cuadrado que hace el algoritmo para generar otra división del árbol, conformando otro nivel en donde la categoría “Rápido” de la siguiente variable independiente que es la de Tempo Musical, vuelve a concentrar la mayoría de sus casos en la categoría “Satisfecho” de la variable dependiente.

Tabla 38 *Matriz de clasificación modo y tempo vs satisfacción*

Pieza	Observado	Pronosticado		
		Insatisfecho	Satisfecho	Porcentaje correcto
Arbol Pieza 1	Insatisfecho	46	16	74%
	Satisfecho	23	51	69%
	Porcentaje global	51%	49%	71%
Arbol Pieza 2	Insatisfecho	75	6	93%
	Satisfecho	38	31	45%
	Porcentaje global	75%	25%	71%
Arbol Pieza 3	Insatisfecho	63	3	95%
	Satisfecho	37	33	47%
	Porcentaje global	74%	26%	71%
Arbol Pieza 4	Insatisfecho	84	10	89%
	Satisfecho	24	23	49%
	Porcentaje global	77%	23%	76%
Arbol Pieza 5	Insatisfecho	64	1	98%
	Satisfecho	42	39	48%
	Porcentaje global	73%	27%	71%

Se observa que en términos generales los porcentajes de clasificación correcta se encuentran alrededor del 71%, lo que puede considerarse un buen porcentaje de clasificación al entender que esto significa que de la muestra aleatoria que se usa para el entrenamiento de cada modelo el 70% de los casos quedan clasificados de forma adecuada es decir que los resultados de los modelos coinciden en un 70% con las verdaderas observaciones. Sin embargo, respecto a la categoría de interés de la variable dependiente que en este caso es la satisfacción (valor 1), se puede apreciar que con excepción de la pieza 1, el porcentaje de clasificación correcto para la categoría satisfecho es menor al 50% en todos los casos, lo que muestra que el buen porcentaje de clasificación correcta de los árboles se debe a la buena clasificación de los casos en la categoría insatisfecho y no a los de la categoría de interés. Solo para la pieza 1 el modelo correspondiente genera un porcentaje relativamente aceptable para estos casos teniendo en cuenta que se acerca al 70%.

Tabla 39 *Riesgo modo y tempo vs satisfacción*

Pieza	Estimación	Tip. Error
P1	,287	,039
P2	,293	,037
P3	,294	,039
P4	,241	,036
P5	,295	,038

La tabla de riesgo muestra en todos los casos estimaciones alrededor del 29 % coherentes con los porcentajes globales de clasificación correcta que en general se sitúan sobre el 71%, se aprecia que la pieza 4 es la que menos riesgo muestra y la que mayor porcentaje de clasificación correcta genera según la tabla 17. Sin embargo, como se mencionaba anteriormente, esto no indica que en su mayoría los modelos generados por los árboles para las diferentes piezas musicales sean buenos para explicar la categoría de interés en la variable independiente “Satisfacción” ya que en su mayoría el porcentaje relacionado con dicha categoría de interés se encuentra por debajo del 50% exceptuando la pieza 1 que genera un porcentaje del 69% en la categoría.

Tabla 40 *Ganancias modo y tempo vs satisfacción*

Pieza	Nodo	Nodo		Ganancia		Respuesta	Índice
		N	Porcentaje	N	Porcentaje		
P1	4	31	23%	30	41%	97%	178%
	3	36	26%	21	28%	58%	107%
	2	69	51%	23	31%	33%	61%
P2	4	37	25%	31	45%	84%	182%
	2	75	50%	27	39%	36%	78%
	3	38	25%	11	16%	29%	63%
P3	3	36	26%	33	47%	92%	178%
	1	65	48%	25	36%	38%	75%
	4	35	26%	12	17%	34%	67%
P4	4	33	23%	23	49%	70%	209%
	2	71	50%	18	38%	25%	76%
	3	37	26%	6	13%	16%	49%
P5	4	40	27%	39	48%	98%	176%
	3	29	20%	12	15%	41%	75%
	2	77	53%	30	37%	39%	70%

La tabla de ganancias muestra que los árboles en general se comportan de acuerdo a lo esperado teniendo en cuenta que los nodos que representan las combinaciones en las que se esperan menos casos en la categoría “Insatisfecho” (0), efectivamente tienen una tasa de respuesta baja y un índice que se ubica debajo de 100 en su mayoría, mientras que los nodos en los que se espera un mayor número de casos ubicados en la categoría “Satisfecho” (1) muestran tasas de respuesta altas e índices mayores y alejados de 100 incluso llegando al 200%, como se observa en el caso de la pieza 4 nodo 4, entendiendo que el Índice representa la tasa de respuesta (porcentaje de casos en la categoría objetivo respecto al total del nodo) relativa al porcentaje de casos en la categoría objetivo del total de la muestra.

- Modo y Tempo VS Número de Parpadeos:

Tabla 41 Árboles CHAID modo y tempo vs número de parpadeos

Arbol	Nodo	Insatisfec		Satisfecho		Total		Categoría pronosticada	Nodo parental	Variable independiente primaria				
		N	%	N	%	N	%			Variable	Sig.a	Chi-cuadrado	gl	Valores de segmentación
Arbol Pieza 1	0	55	42%	76	58%	131	100%	Hartos						
	1	31	42%	43	58%	74	56%	Hartos	0	Modo Musical Dicotómico	0,98	0	1	Modo Mayor
	2	24	42%	33	58%	57	44%	Hartos	0	Modo Musical Dicotómico	0,98	0	1	Modo Menor
	3	22	58%	16	42%	38	29%	Pocos	1	Tempo Musical Dicotómica	0,00	8	1	Lento
	4	9	25%	27	75%	36	27%	Hartos	1	Tempo Musical Dicotómica	0,00	8	1	Rápido
	5	16	64%	9	36%	25	19%	Pocos	2	Tempo Musical Dicotómica	0,00	9	1	Lento
Arbol Pieza 2	6	8	25%	24	75%	32	24%	Hartos	2	Tempo Musical Dicotómica	0,00	9	1	Rápido
	0	79	56%	62	44%	141	100%	Pocos						
	1	31	45%	38	55%	69	49%	Hartos	0	Modo Musical Dicotómico	0,01	7	1	Modo Mayor
	2	48	67%	24	33%	72	51%	Pocos	0	Modo Musical Dicotómico	0,01	7	1	Modo Menor
	3	21	66%	11	34%	32	23%	Pocos	1	Tempo Musical Dicotómica	0,00	11	1	Lento
	4	10	27%	27	73%	37	26%	Hartos	1	Tempo Musical Dicotómica	0,00	11	1	Rápido
Arbol Pieza 3	5	32	84%	6	16%	38	27%	Pocos	2	Tempo Musical Dicotómica	0,00	11	1	Lento
	6	16	47%	18	53%	34	24%	Hartos	2	Tempo Musical Dicotómica	0,00	11	1	Rápido
	0	85	63%	49	37%	134	100%	Pocos						
	1	36	53%	32	47%	68	51%	Pocos	0	Modo Musical Dicotómico	0,01	7	1	Modo Mayor
	2	49	74%	17	26%	66	49%	Pocos	0	Modo Musical Dicotómico	0,01	7	1	Modo Menor
	3	9	28%	23	72%	32	24%	Hartos	1	Tempo Musical Dicotómica	0,00	16	1	Rápido
Arbol Pieza 4	4	27	75%	9	25%	36	27%	Pocos	1	Tempo Musical Dicotómica	0,00	16	1	Lento
	5	18	51%	17	49%	35	26%	Pocos	2	Tempo Musical Dicotómica	0,00	27	1	Rápido
	6	31	100%	0	0%	31	23%	Pocos	2	Tempo Musical Dicotómica	0,00	27	1	Lento
	0	86	61%	56	39%	142	100%	Pocos						
	1	38	50%	38	50%	76	54%	Pocos	0	Modo Musical Dicotómico	0,01	8	1	Modo Mayor
	2	48	73%	18	27%	66	46%	Pocos	0	Modo Musical Dicotómico	0,01	8	1	Modo Menor
Arbol Pieza 5	3	30	81%	7	19%	37	26%	Pocos	1	Tempo Musical Dicotómica	0,00	30	1	Lento
	4	8	21%	31	79%	39	27%	Hartos	1	Tempo Musical Dicotómica	0,00	30	1	Rápido
	0	91	65%	48	35%	139	100%	Pocos						
	1	38	56%	30	44%	68	49%	Pocos	0	Modo Musical Dicotómico	0,02	5	1	Modo Mayor
Arbol Pieza 5	2	53	75%	18	25%	71	51%	Pocos	0	Modo Musical Dicotómico	0,02	5	1	Modo Menor
	3	31	89%	4	11%	35	25%	Pocos	1	Tempo Musical Dicotómica	0,00	34	1	Lento
	4	7	21%	26	79%	33	24%	Hartos	1	Tempo Musical Dicotómica	0,00	34	1	Rápido

A diferencia de los árboles generados en el análisis anterior, para este caso en que el Número de parpadeos binario es la variable dependiente, en las piezas musicales 1,2 y 3 el árbol llega a desarrollar todas sus ramas por completo creando así 6 nodos mientras que para las piezas 4 y 5 no sucede lo mismo.

De igual forma que en el análisis anterior se observa que la variable modo musical es la que conforma el segundo nivel del árbol en todas las piezas musicales, sugiriendo así su mayor relación explicativa con la variable dependiente

De otro lado se observa que en la pieza 2 es en la única en donde el árbol muestra en todos sus niveles y nodos el comportamiento esperado ya que en el segundo nivel conformado por la variable Modo Musical la mayoría de casos para el modo Mayor corresponden con el valor de la variable independiente “Hartos” (1) y para el Modo menor la mayoría de casos corresponden con el valor “Pocos” (0) de la variable dependiente.

Tabla 42 *Matriz de clasificación modo y tempo vs número de parpadeos*

Pieza	Observado	Pronosticado		
		Insatisfecho	Satisfecho	Porcentaje correcto
Arbol Pieza 1	Pocos	38	17	69%
	Hartos	25	51	67%
	Porcentaje global	48%	52%	68%
Arbol Pieza 2	Pocos	53	26	67%
	Hartos	17	45	73%
	Porcentaje global	50%	50%	70%
Arbol Pieza 3	Pocos	76	9	89%
	Hartos	26	23	47%
	Porcentaje global	76%	24%	74%
Arbol Pieza 4	Pocos	78	8	91%
	Hartos	25	31	55%
	Porcentaje global	73%	27%	77%
Arbol Pieza 5	Pocos	84	7	92%
	Hartos	22	26	54%
	Porcentaje global	76%	24%	79%

En la matriz de clasificación para la variable explicada Número de Parpadeos, los porcentajes globales de clasificación correcta son en términos generales mayores que en el caso anterior, esto teniendo en cuenta que se ubican en general por encima del 70 % en la mayoría de los casos. Respecto al porcentaje de clasificación correcta de la categoría objetivo “Hartos”, se puede decir que también es en general mayor respecto al caso anterior y en dos de las piezas (1 y 2) es cercano al 70%, sin embargo, en las otras 3 (piezas 3, 4 y 5) su comportamiento dista del esperado ya que tiene valores que se sitúan cerca al 50%, lo que no representa modelos confiables de predicción en estos casos.

Tabla 43 *Riesgo modo y tempo vs número de parpadeos*

Pieza	Estimación	Tip. Error
P1	,321	,041
P2	,305	,039
P3	,261	,038
P4	,232	,035
P5	,209	,034

La tabla de riesgo para este caso genera resultados de riesgo para los árboles ligeramente distintos a los del caso anterior si se observa que los porcentajes de estimación del riesgo para las piezas 1 y 2 corresponden a valores cercanos al 30%, mientras que para las demás piezas musicales se sitúan entre el 20% y 26% coherente con los porcentajes globales de clasificación, sin embargo, aunque los valores de riesgo estimado de las dos primeras piezas

(1 y 2) son ligeramente mayores que en las otras piezas, puede decirse que son mejores modelos ya que el porcentaje de clasificación correcta de la categoría objetivo para estas 2 piezas es significativamente mayor que el de las otras 3.

Tabla 44 *Ganancias modo y tempo VS número de parpadeos*

Pieza	Nodo	Nodo		Ganancia		Respuesta	Índice
		N	Porcentaje	N	Porcentaje		
P1	4	36	27%	27	36%	75%	129%
	6	32	24%	24	32%	75%	129%
	3	38	29%	16	21%	42%	73%
	5	25	19%	9	12%	36%	62%
P2	4	37	26%	27	44%	73%	166%
	6	34	24%	18	29%	53%	120%
	3	32	23%	11	18%	34%	78%
	5	38	27%	6	10%	16%	36%
P3	3	32	24%	23	47%	72%	197%
	5	35	26%	17	35%	49%	133%
	4	36	27%	9	18%	25%	68%
	6	31	23%	0	0%	0%	0%
P4	4	39	27%	31	55%	79%	202%
	2	66	46%	18	32%	27%	69%
	3	37	26%	7	13%	19%	48%
P5	4	33	24%	26	54%	79%	228%
	2	71	51%	18	38%	25%	73%
	3	35	25%	4	8%	11%	33%

Al igual que en el caso anterior, la tabla muestra que los nodos reflejan el comportamiento esperado teniendo en cuenta que el índice de la tasa de respuesta es mayor a 100 en los nodos en los que se espera que haya más casos con respuesta “Hartos” y que corresponden a las combinaciones de Modo mayor y Tempo rápido, mientras que para los demás casos en los que se espera una menor concentración de casos de la categoría “Hartos” el Índice es menor a 100 y se aleja bastante de este valor, estos corresponden a los nodos que representan las combinaciones de Modo Menor y Tempo Lento. Se observan índices bastante fuertes en el nodo 4 de las piezas musicales 4 y 5 teniendo en cuenta que el porcentaje de respuesta es bastante alto respecto al porcentaje de casos en la categoría “Hartos” en el total de la muestra, en donde la relación es de más del doble.

- Modo y Tempo VS Tiempo de Respuesta:

Tabla 45 Árboles CHAID modo y tempo vs tiempo de respuesta

Arbol	Nodo	Insatisfec		Satisfecho		Total		Categoría pronosticada	Nodo parental	Variable independiente primaria				
		N	%	N	%	N	%			Variable	Sig.a	Chi-cuadrado	gl	Valores de segmentación
Arbol Pieza 1	0	73	57%	56	43%	129	100%	Lento						
	1	45	64%	25	36%	70	54%	Lento	0	Modo Musical Dicotómico	0,05	4	1	Modo Menor
	2	28	47%	31	53%	59	46%	Rápido	0	Modo Musical Dicotómico	0,05	4	1	Modo Mayor
	3	21	72%	8	28%	29	22%	Lento	2	Tempo Musical (velocidad) Dicotómico	0,00	15	1	Lento
	4	7	23%	23	77%	30	23%	Rápido	2	Tempo Musical (velocidad) Dicotómico	0,00	15	1	Rápido
Arbol Pieza 2	0	81	57%	62	43%	143	100%	Lento						
	1	50	68%	23	32%	73	51%	Lento	0	Modo Musical Dicotómico	0,00	9	1	Modo Menor
	2	31	44%	39	56%	70	49%	Rápido	0	Modo Musical Dicotómico	0,00	9	1	Modo Mayor
	3	8	24%	26	76%	34	24%	Rápido	2	Tempo Musical Dicotómica	0,00	12	1	Rápido
	4	23	64%	13	36%	36	25%	Lento	2	Tempo Musical Dicotómica	0,00	12	1	Lento
Arbol Pieza 3	0	87	58%	63	42%	150	100%	Lento						
	1	52	68%	24	32%	76	51%	Lento	0	Modo Musical Dicotómico	0,01	7	1	Modo Menor
	2	35	47%	39	53%	74	49%	Rápido	0	Modo Musical Dicotómico	0,01	7	1	Modo Mayor
	3	7	19%	30	81%	37	25%	Rápido	2	Tempo Musical Dicotómica	0,00	25	1	Rápido
	4	28	76%	9	24%	37	25%	Lento	2	Tempo Musical Dicotómica	0,00	25	1	Lento
Arbol Pieza 4	0	75	59%	52	41%	127	100%	Lento						
	1	27	43%	36	57%	63	50%	Rápido	0	Modo Musical Dicotómico	0,00	14	1	Modo Mayor
	2	48	75%	16	25%	64	50%	Lento	0	Modo Musical Dicotómico	0,00	14	1	Modo Menor
	3	23	68%	11	32%	34	27%	Lento	1	Tempo Musical Dicotómica	0,00	20	1	Lento
	4	4	14%	25	86%	29	23%	Rápido	1	Tempo Musical Dicotómica	0,00	20	1	Rápido
Arbol Pieza 5	0	77	55%	63	45%	140	100%	Lento						
	1	45	62%	28	38%	73	52%	Lento	0	Modo Musical Dicotómico	0,10	3	1	Modo Menor
	2	32	48%	35	52%	67	48%	Rápido	0	Modo Musical Dicotómico	0,10	3	1	Modo Mayor
	3	22	65%	12	35%	34	24%	Lento	2	Tempo Musical Dicotómica	0,00	8	1	Lento
	4	10	30%	23	70%	33	24%	Rápido	2	Tempo Musical Dicotómica	0,00	8	1	Rápido

Para los árboles corridos para la variable Tiempo de Respuesta dicotómico se aprecia que para todas las piezas musicales el nodo inicial 0 los casos para las muestras de entrenamiento que corresponden a los casos “Rápido” son menos al representar entre un 43% y 45%, sin embargo, para todas las piezas musicales nuevamente la variable más significativa al definir la categoría objetivo que en este caso es el tiempo de Respuesta “Rápido”, es el Modo y los casos se distribuyen también según lo esperado ya que para todas las 5 piezas musicales la mayoría de los casos se agrupan en el Modo Mayor. En el tercer nivel del árbol de igual forma los casos también se agrupan en su mayoría en el Tempo Rápido para esta variable en concordancia con lo que se espera para la categoría objetivo Tiempo de Respuesta “Rápido”.

Tabla 46 *Matriz de clasificación modo y tempo vs tiempo de respuesta*

Pieza	Observado	Pronosticado		
		Insatisfecho	Satisfecho	Porcentaje correcto
Arbol Pieza 1	Lento	66	7	90%
	Rápido	33	23	41%
	Porcentaje global	77%	23%	69%
Arbol Pieza 2	Lento	73	8	90%
	Rápido	36	26	42%
	Porcentaje global	76%	24%	69%
Arbol Pieza 3	Lento	80	7	92%
	Rápido	33	30	48%
	Porcentaje global	75%	25%	73%
Arbol Pieza 4	Lento	71	4	95%
	Rápido	27	25	48%
	Porcentaje global	77%	23%	76%
Arbol Pieza 5	Lento	67	10	87%
	Rápido	40	23	37%
	Porcentaje global	76%	24%	64%

En esta matriz de clasificación se aprecian porcentajes de clasificación correcta cercanos al 70 % para las piezas 1, 2 y 3, mientras que en las piezas 4 y 5 se encuentran el mayor y menor valor de los mismos con un 76% y 64% respectivamente. Respecto al porcentaje de clasificación correcta para la categoría objetivo que en este caso es “Rápido” en todos los casos es de los más bajos que se han visto hasta el momento al estar alrededor del 48% o menos en todos los casos. Para los árboles generados con esta variable explicada Tiempo de respuesta no hay ningún modelo con valores aceptables respecto a porcentajes de clasificación correcta para los valores de interés de la variable objetivo.

Tabla 47 *Riesgo modo y tempo vs tiempo de respuesta*

Pieza	Estimación	Tip. Error
P1	,310	,041
P2	,308	,039
P3	,267	,036
P4	,244	,038
P5	,357	,040

Respecto a las estimaciones de riesgo de los modelos aquí generados, se puede ver que son similares a los de los dos anteriores casos solo valores ligeramente mayores situadas cerca

al 30% en la mayoría de los casos, sin embargo, aunque estos modelos tienen porcentajes de estimación de riesgo similares, los de clasificación correcta para las variables objetivo si difieren ya que en este caso ninguno resulta aceptable mientras que en los anteriores casos esto si sucede con los árboles generados para las piezas 1 en el primer caso (variable respuesta Satisfacción) y 1 y 2 para el segundo (Variable respuesta Número de Parpadeos).

Tabla 48 *Ganancias modo y tempo vs tiempo de respuesta.*

Pieza	Nodo	Nodo		Ganancia		Respuesta	Índice
		N	Porcentaje	N	Porcentaje		
P1	4	30	23%	23	41%	77%	177%
	1	70	54%	25	45%	36%	82%
	3	29	22%	8	14%	28%	64%
P2	3	34	24%	26	42%	76%	176%
	4	36	25%	13	21%	36%	83%
	1	73	51%	23	37%	32%	73%
P3	3	37	25%	30	48%	81%	193%
	1	76	51%	24	38%	32%	75%
	4	37	25%	9	14%	24%	58%
P4	4	29	23%	25	48%	86%	211%
	3	34	27%	11	21%	32%	79%
	2	64	50%	16	31%	25%	61%
P5	4	33	24%	23	37%	70%	155%
	1	73	52%	28	44%	38%	85%
	3	34	24%	12	19%	35%	78%

Se observan Índices mayores a 100 en los casos en que se espera una mayor proporción de casos para la categoría “Rápido” de la variable Objetivo, lo contrario sucede con los nodos en los que se espera menor cantidad de casos “Rápido”, lo que refleja un comportamiento adecuado de los árboles desde este punto de vista, no obstante, los bajos porcentajes de clasificación correcta de los casos de la variable objetivo no permiten la categorización de estos árboles como modelos confiables.

- Modo y Tempo VS Latencia:

Tabla 49 Árboles CHAID modo y tempo vs latencia

Arbol	Nodo	Insatisfec		Satisfecho		Total		Categoría pronosticada	Nodo parental	Variable independiente primaria				
		N	%	N	%	N	%			Variable	Sig.a	Chi-cuadrado	Valores de segmentación	
Arbol Pieza 1	0	83	59%	57	41%	140	100%	Poca						
	1	37	50%	37	50%	74	53%	Poca	0	Modo Musical Dicotómico	0,02	6	1	Modo Mayor
	2	46	70%	20	30%	66	47%	Poca	0	Modo Musical Dicotómico	0,02	6	1	Modo Menor
	3	28	72%	11	28%	39	28%	Poca	1	Tempo Musical Dicotómica	0,00	16	1	Lento
Arbol Pieza 2	4	9	26%	26	74%	35	25%	Harto	1	Tempo Musical Dicotómica	0,00	16	1	Rápido
	0	76	54%	65	46%	141	100%	Poca						
	1	31	42%	42	58%	73	52%	Harto	0	Modo Musical Dicotómico	0,00	8	1	Modo Mayor
	2	45	66%	23	34%	68	48%	Poca	0	Modo Musical Dicotómico	0,00	8	1	Modo Menor
Arbol Pieza 3	3	25	64%	14	36%	39	28%	Poca	1	Tempo Musical Dicotómica	0,00	17	1	Lento
	4	6	18%	28	82%	34	24%	Harto	1	Tempo Musical Dicotómica	0,00	17	1	Rápido
	0	88	61%	57	39%	145	100%	Poca						
	1	40	52%	37	48%	77	53%	Poca	0	Modo Musical Dicotómico	0,02	5	1	Modo Mayor
Arbol Pieza 4	2	48	71%	20	29%	68	47%	Poca	0	Modo Musical Dicotómico	0,02	5	1	Modo Menor
	3	32	78%	9	22%	41	28%	Poca	1	Tempo Musical Dicotómica	0,00	25	1	Lento
	4	8	22%	28	78%	36	25%	Harto	1	Tempo Musical Dicotómica	0,00	25	1	Rápido
	0	80	61%	51	39%	131	100%	Poca						
Arbol Pieza 5	1	31	48%	33	52%	64	49%	Harto	0	Modo Musical Dicotómico	0,00	8	1	Modo Mayor
	2	49	73%	18	27%	67	51%	Poca	0	Modo Musical Dicotómico	0,00	8	1	Modo Menor
	3	8	26%	23	74%	31	24%	Harto	1	Tempo Musical Dicotómica	0,00	13	1	Rápido
	4	23	70%	10	30%	33	25%	Poca	1	Tempo Musical Dicotómica	0,00	13	1	Lento
Arbol Pieza 5	0	77	59%	54	41%	131	100%	Poca						
	1	34	52%	32	48%	66	50%	Poca	0	Modo Musical Dicotómico	0,09	3	1	Modo Mayor
	2	43	66%	22	34%	65	50%	Poca	0	Modo Musical Dicotómico	0,09	3	1	Modo Menor
	3	23	72%	9	28%	32	24%	Poca	1	Tempo Musical Dicotómica	0,00	11	1	Lento
Arbol Pieza 5	4	11	32%	23	68%	34	26%	Harto	1	Tempo Musical Dicotómica	0,00	11	1	Rápido

En el caso de la variable dependiente Latencia, para todos los casos nuevamente resulta la variable Modo como la más significativa para explicar la categoría Objetivo “Harta”. Sin embargo, los resultados generados por los árboles no son tan correspondientes con lo esperado en términos generales ya que, en 3 de las piezas musicales, la 1, la 3 y la 5, El Modo Musical no tiene la mayoría de sus casos con respuesta “Harto” en su categoría “Mayor”. Respecto a la variable Tempo Musical se aprecia, por el contrario, que si presenta el comportamiento esperado ya que para todas las piezas musicales la mayor parte de los casos con respuesta Harto se encuentran en la categoría “Rápido” y la menor en la categoría “Lento”.

Tabla 50 *Matriz de clasificación modo y tempo vs latencia*

Pieza	Observado	Pronosticado		
		Insatisfecho	Satisfecho	Porcentaje correcto
Arbol Pieza 1	Poca	74	9	89%
	Harto	31	26	46%
	Porcentaje global	75%	25%	71%
Arbol Pieza 2	Poca	70	6	92%
	Harto	37	28	43%
	Porcentaje global	76%	24%	70%
Arbol Pieza 3	Poca	80	8	91%
	Harto	29	28	49%
	Porcentaje global	75%	25%	74%
Arbol Pieza 4	Poca	72	8	90%
	Harto	28	23	45%
	Porcentaje global	76%	24%	73%
Arbol Pieza 5	Poca	66	11	86%
	Harto	31	23	43%
	Porcentaje global	74%	26%	68%

Finalmente se tiene la matriz de clasificación para la variable respuesta dicotomizada “Latencia” o duración del parpadeo, se aprecian porcentajes globales de clasificación correcta que en general están por encima del 70%, sin embargo, respecto a lo esperado, estos porcentajes son bastante bajos teniendo en cuenta que en todos los casos están por debajo del 50% para la categoría objetivo “Harto” (1).

Tabla 51 *Riesgo modo y tempo vs latencia*

Pieza	Estimación	Típ. Error
P1	,286	,038
P2	,305	,039
P3	,255	,036
P4	,275	,039
P5	,321	,041

Respecto al riesgo se aprecian valores promedio de estimaciones de riesgo del 28,6%, lo que en términos generales ubica los resultados de estos árboles en un segundo lugar respecto a las estimaciones de riesgo más altas de las 4 variables respuesta analizadas, siendo la pieza 5 la del porcentaje más alto con un 30,5% y la 3 con el más bajo (25,5%).

Tabla 52 *Ganancias modo y tempo vs latencia*

Pieza	Nodo	Nodo		Ganancia		Respuesta	Índice
		N	Porcentaje	N	Porcentaje		
P1	4	35	25%	26	46%	74%	182%
	2	66	47%	20	35%	30%	74%
	3	39	28%	11	19%	28%	69%
P2	4	34	24%	28	43%	82%	179%
	3	39	28%	14	22%	36%	78%
	2	68	48%	23	35%	34%	73%
P3	4	36	25%	28	49%	78%	198%
	2	68	47%	20	35%	29%	75%
	3	41	28%	9	16%	22%	56%
P4	3	31	24%	23	45%	74%	191%
	4	33	25%	10	20%	30%	78%
	2	67	51%	18	35%	27%	69%
P5	4	34	26%	23	43%	68%	164%
	2	65	50%	22	41%	34%	82%
	3	32	24%	9	17%	28%	68%

Respecto a la tabla de ganancias los índices del nodo 4 de los árboles se encuentran todos por encima del 100% indicando que efectivamente el porcentaje de casos de la categoría “Rápido” de la variable Tempo y que también pertenecen a la categoría “Mayor” de la variable Modo y a la categoría “Harto” de la variable respuesta, siempre es mayor que los que pertenecen a las mismas categorías de las variables explicativas pero que corresponden a la categoría “Poco” de dicha variable. Es decir que en términos generales la tabla de ganancia muestra que los árboles se comportan según lo esperado, no obstante, la capacidad predictiva de los árboles para la variable Latencia no es confiable por sus bajos porcentajes de clasificación correcta en la categoría objetivo “harto”.

Conclusiones de los árboles CHAID.

Viendo de modo transversal los comportamientos que tuvieron los árboles CHAID desarrollados para las 4 variables respuesta y para cada pieza musical, se puede concluir que de forma general se presentan los comportamientos esperados en todas las piezas musicales para todas las variables respuesta analizadas, ya que en todos los casos las tablas de ganancia muestran índices mayores a 100 en los nodos terminales donde el porcentaje de casos que

pertenece a las categorías objetivo de las diferentes variables respuesta es mayor para las categorías objetivo correspondientes que para la otra categoría, esto después de pertenecer a las categorías 1 de las variables independientes Modo y Tempo, lo que nos indicaría la presencia de argumentos estadísticos para pensar en que efectivamente hay una relación de causalidad entre el Modo Mayor y el Tempo Rápido con la presencia de Satisfacción en las diferentes Variables Explicadas. (Satisfacción, Tiempo de Respuesta, Latencia y Número de parpadeos dicotómicos).

Otro de los aspectos importantes que se observa en todos los casos es que la variable Modo Musical siempre aparece en el segundo nivel de los árboles, lo que indica que es la que tiene el mayor impacto sobre la categoría objetivo de las diferentes variables respuesta de acuerdo a las pruebas Ji Cuadrado que utiliza el algoritmo CHAID por lo que su poder explicativo sobre los resultados de las diferentes variables respuesta es estadísticamente más significativo que el de la Variable Tempo Musical que en todos los casos se encuentra en el último nivel de los árboles.

De otro lado, los árboles para todas las piezas en todas las variables respuesta muestran porcentajes de clasificación correcta mayores al 70% en su mayoría lo que indicaría a primera vista un buen poder de clasificación de los modelos creados con esas muestras de entrenamiento aleatorias (70% del total de los casos). Sin embargo, al profundizar en la composición de dichos porcentajes de clasificación correcta, en casi la totalidad de los casos, este corresponde a la categoría 0, de las variables respuesta, la cual no es la objetivo, por el contrario los porcentajes de clasificación correcta de los casos que pertenecen a las categorías objetivo de las variables respuesta (1), son en su mayoría cercanos al 50%, sugiriendo así que los modelos clasifican mejor los casos de las categorías 0 que los de las 1 en las diferentes variables respuesta analizadas. Sin embargo, es importante resaltar que para los árboles creados para la variable respuesta Satisfacción Dicotómica en la pieza 1 y para los de las piezas 1 y 2

para la variable respuesta Número de Parpadeos, los modelos de clasificación si son robustos ya que sus porcentajes de clasificación correcto para los casos con valor 1 de las respectivas variables objetivo se encuentran alrededor del 70%.

Respecto al Riesgo estimado no hay mayores diferencias en términos generales si se tiene en cuenta que en todos los casos las estimaciones se encuentran cercanas al 28% lo que resulta consistente con niveles globales de clasificación cercanos al 70%, sin embargo, como se mencionó anteriormente, dichos porcentajes aunque son aceptables a nivel global no lo son para las categorías objetivo con excepción de los modelos de las piezas 1 y 1 y 2 creados para las variables respuesta Satisfacción Dicotómica y Número de parpadeos Dicotómico, en cuyos casos la evidencia estadística apuntaría a los resultados esperados en todo sentido.

Finalmente y de acuerdo a todo lo anterior, se puede concluir que aunque los árboles de clasificación creados para las distintas piezas musicales y las diferentes variables respuesta se comportan de acuerdo a lo esperado (como lo muestran las tablas de ganancia y las de los mismos árboles), en términos de las respectivas categorías objetivo, solo son confiables en su totalidad para los casos de las piezas 1 y 1 y 2 en las variables respuesta Satisfacción y Número de parpadeos dicotómicas ya que es en éstos donde los porcentajes de clasificación correcta para las categorías objetivo es aceptable (alrededor del 70%) además de mostrar porcentajes de estimación de riesgos aceptables y comportamientos acordes a lo esperado según los resultados de las categorías objetivo y sus valores en las categorías de las variables explicativas en donde el Modo Mayor (1) y el Tempo Rápido (1) generan respuestas de Satisfecho (1) y Hartos (1) para el primer y segundo caso respectivamente.

Conclusiones Generales Del Análisis.

De acuerdo a lo que se puede ver a lo largo del análisis realizado en el presente trabajo, es notable desde el principio, la presencia de dependencia existente entre las variables independientes y las explicativas.

Empezando con el análisis exploratorio en un inicio el univariado ya estaba sugiriendo tendencias de concentración definidas. Como ya se vio en el apartado correspondiente, se observan coeficientes de variación altos y además asimetrías positivas para todas las variables dependientes, lo que en una primera instancia sugiere resultados polarizados que posiblemente sean generados por la presencia o el efecto de las variables independientes como se comprueba más adelante en el correspondiente análisis bivariado.

De este modo, según lo observado en el correspondiente análisis bivariado, hay evidencia estadística para sugerir la presencia de dependencia entre las variables explicadas y las independientes, teniendo en cuenta que en todos los casos, con excepción de uno (Variable Número de Parpadeos Dicotómica, pieza 1 variable independiente Modo Musical), las pruebas chi cuadrado resultan significativas con p valores menores a 0,05, rechazando en todos los casos la Hipótesis Nula de que no Hay dependencia entre las variables explicativas y las dependientes. Finalmente, no solo se observa que las pruebas de dependencia son significativas, sino que además de esto los estadísticos que miden la intensidad de la relación muestran valores importantes que en donde se concluyen relaciones en general fuertes entre las variables independientes y las explicadas.

Respecto a las piezas musicales, se ha observado que para el análisis exploratorio en general, esto es, univariado y bivariado, son las piezas 1,4 y la 5 las que muestran comportamientos sobresalientes y marcados respecto a las demás, especialmente la 1 que se comporta según lo esperado con tendencias consistentes y marcadas en términos de satisfacción, ya que en el análisis bivariado los casos de satisfacción son mayores que en las demás cuando las respectivas variables independientes toman el valor con el que se espera haya satisfacción: Modo Mayor y Tempo Rápido en las variables Satisfacción dicotómica y Tiempo de Respuesta.

Ya el encontrar que existe evidencia de la relación de dependencia entre cada variable explicada con sus dependientes según los resultados anteriores, se aplican las técnicas que más se ajustan a la naturaleza de los datos, con las visiones del punto de vista paramétrico como no paramétrico. Así se aplican pruebas de Wicoxon en un primer lugar desde el punto de vista no paramétrico, y en estas se concluye que nuevamente hay una relación de dependencia entre las variables respuesta y las independientes ya que en todos los casos las pruebas son significativas con valores p menores a 0,05 rechazando las Hipótesis nulas de que no hay diferencias en los valores que toman las variables dependientes en la presencia de cada una de las dependientes. Desde el punto de vista paramétrico se aplican regresiones logísticas y árboles de decisión en un segundo y tercer lugar. De acuerdo con esto en las regresiones logísticas se observa nuevamente significancia respecto a las relaciones de las variables dependientes y las explicativas ya que los estadísticos de Wald generan valores p menores a 0.05 en todos los casos, en cuyos casos implica que siempre hay efectos de las variables independientes sobre las explicadas. Respecto a la bondad de ajuste de las regresiones generadas para todas estas variables en las diferentes piezas musicales, se observan las diferentes tablas en donde están los Pseudo R cuadrados de Cos y Snell y sus correspondientes ajustes con los de Nagelkerke, desde un punto de vista no tan riguroso pareciera que en términos generales los modelos explican la variabilidad en las diferentes variables explicadas y para cada pieza musical, sin embargo desde un punto de vista más riguroso las Pruebas de Hosmer y Lemeshov que complementan este análisis de Bondad de Ajuste salen significativas lo que implica el rechazo de la Hipótesis Nula de que No hay diferencias entre los valores pronosticados y los reales, lo que implica que los modelos no tienen buena capacidad predictiva desde este punto de vista más riguroso.

En cuanto a la evaluación de los respectivos modelos, se observan porcentajes de clasificación correcta globales que en general se encuentran el rededor del 70%, sin embargo,

en la mayoría de los casos este no se encuentra compuesto por el porcentaje de clasificación correcta de la categoría objetivo, sino de la de que representa insatisfacción, esto en todas las variables menos en la de Satisfacción dicotómica que en su mayoría si cuenta con un porcentaje de clasificación correcta compuesto por los casos de la categoría objetivo.

De otro lado, es importante ver que los ODD ratio (exp B) de todas las regresiones son menores a uno, lo que implica que los resultados de predicción no se comportan según lo esperado en donde las categorías de las variables independientes con valor 1 (modo, Mayor y Tempo Rápido) no generan una mayor probabilidad de ocurrencia de Satisfacción (1) en las variables explicadas sino por el contrario de que el valor sea de insatisfacción (0).

Finalmente, respecto a las piezas musicales, se vuelve a presentar comportamientos sobresalientes de las piezas 1 y 4 en las variables en las variables de Satisfacción Dicotómica y Latencia del parpadeo Dicotómica) en donde los porcentajes de clasificación correcta de dichas variables a nivel global son mayores al de las demás, hay que resaltar que en este caso se trata de Porcentajes de clasificación correcta en las categorías objetivo, para estas dos variables. La pieza 3 presenta comportamiento sobresaliente en este sentido en la variable # de Parpadeos dicotómica en donde el mayor porcentaje de clasificación correcta se da por la categoría "Pocos", en cuanto a la pieza 5 en general presenta los indicadores más distantes de lo esperado ya que en todas las regresiones se localiza en el porcentaje de clasificación correcto más bajo, en cuanto a la pieza 2 no se observa ningún comportamiento diferencial particular.

En cuanto al análisis con los árboles de decisión, en una primera instancia se observa que se presentan los comportamientos esperados en las variables objetivo analizadas, ya que las tablas de ganancia muestran en todos los casos porcentajes mayores a 100 en los índices de los nodos terminales lo que indica que los casos acumulados en la categoría objetivo siempre son mayores respecto a la otra.

De otro lado es importante observar que la variable Modo Musical es la que siempre aparece en el segundo nodo de todos los árboles lo que indica que es la que más influencia tiene sobre las diferentes variables respuesta analizadas, teniendo en cuenta que el algoritmo del árbol CHAID realiza pruebas Chi cuadrado para elegir la variable en la que dividirá el árbol según su poder explicativo sobre la independiente.

Al igual que en las regresiones, las matrices de clasificación de estos árboles muestran porcentajes de clasificación correcta que en su mayoría se encuentran sobre el 70%, sin embargo, dicho porcentaje está compuesto, en casi todos los casos, por la clasificación correcta del valor 0 y solo en los árboles correspondientes a la pieza 1 y la variable Satisfacción dicotómica y a las piezas 1 y 2 y la variable Número de Parpadeos dicotómica este porcentaje lo representa en mayor medida la categoría objetivo correspondiente. Consistente con lo anterior se comportan los niveles de riesgo estimado que se encuentran en la mayor parte de los casos alrededor del 30%.

Como se ha podido observar es concluyente que efectivamente se presenta dependencia entre las variables explicativas Modo y Tempo Musical, y las variables dicotimizadas Satisfacción, Número de Parpadeos, Tiempo de Respuesta Y Latencia, esto teniendo en cuenta que en las 3 técnicas utilizadas se presenta significancia en este sentido. De otro lado transversalmente en los diferentes análisis realizados también se observan comportamientos estables de unas de las piezas musicales respecto a las demás, así se puede observar que en especial la pieza 1 que muestra tener el comportamiento esperado respecto a la variable satisfacción dicotómica en donde los porcentajes de clasificación correcta tanto en la regresión logística como en los árboles de decisión son mayores en la categoría objetivo 1 “Satisfecho” que en la 0 “Insatisfecho”, agregando a esto que también las pruebas correspondientes de dependencia dieron significativas incluyendo las pruebas de Wilcoxon, de forma contraria la pieza 5 muestra tener el comportamiento inverso al esperado con más fuerza que las demás

piezas musicales ya que aunque las pruebas de dependencia dan significativas, los porcentajes de clasificación correcta se concentran en la categoría 0 tanto en la variable Satisfacción dicotómica como en las demás.

Referencias

- Abgar, N., & Zuniga, C. (2006). *Breve Aproximación la Técnica de Árbol de Decisiones.*, 2006.
- Abuin, J. R., & Rojo, J. M. (2007) Regresión con variable dependiente cualitativa.
- Araya, M. & Magno. G. (1994). *Segmentación de Mercados usando la técnica CHAID*, Universidad de Costa Rica
- Baños, R. & García P. (1999). *Eficacia de dos procedimientos de inducción del estado de ánimo e influencia de variables moduladoras. Revista de Psicopatología y Psicología Clínica. 4(1) 15-26.*
- Barreiro, J. Ruzo, E. & Losada, F. (2004) Modelo Logit Multinomial: Una Aplicación al sector Lácteo, Regional and Sectoral Economic Studies
- Benedetti F. (SA) . Efectos de Tempo y Modo sobre el contenido emocional de la música. Universitat Jaume.
- Berlanga-Silvente, V., Rubio-Hurtado, M. J., & Baños, R. V. (2013). Como aplicar Árboles de decisión en SPSS. *REIRE Revista d'Innovació i Recerca en Educació*, 6(1), 65-79.
- Caballero Meneses A., J., & Menez, M. (2010). Influencia del tempo de la música en las emociones. *Revista Colombiana de Psicología*, 19(1), 37-44.
- Caldwell, C., & Hibbert, S. A. (2002). The influence of music tempo and musical preference on restaurant patrons' behavior. *Psychology & Marketing*, 19(11), 895-917.
- Chambers, C. P., Echenique, F., & Shmaya, E. (2017). General revealed preference theory. *Theoretical Economics*, 12(2), 493-511.
- Croissant I. (2010), Estimation of multinomial models in R: The mlogit Packages. Recuperado de <https://cran.r-project.org/web/packages/mlogit/vignettes/mlogit.pdf>

Del Rio L J. (s.f). Introducción a SPSS Árboles de clasificación recuperado de <http://revistes.ub.edu/index.php/REIRE/article/viewFile/reire2013.6.1615/7229>

Díaz, L. G., & Morales, M. (2002). Análisis estadístico de datos categóricos. *Notas de Clase del Departamento de Estadística de la Universidad Nacional de Colombia. Bogotá: Universidad Nacional de Colombia.*

Domjan, M. (2007). *Principios de aprendizaje y conducta*. Editorial Paraninfo.

Doughty, M. J. (2002). Further assessment of gender-and blink pattern-related differences in the spontaneous eyeblink activity in primary gaze in young adult humans. *Optometry & Vision Science*, 79(7), 439-447.

Fernández-Collado, H., Sampieri Hernández R. & Baptista Lucio P (2006). *Metodología De La Investigación* (No. H61. H47 2006.).

Fierro, J. (2010). Análisis estadístico univariado, bivariado y variables control. *Material docente, curso "Análisis de Datos Cuantitativos". Depto. Sociología, Facultad de Ciencias Sociales, Universidad de Chile.*

Litherland N.(s.f). Que tan rápido es el parpadeo de un ojo?. Recuperado de http://www.ehowenespanol.com/rapido-parpadeo-ojo-sobre_155659/.

Mazzocchi, M. (2008). *Statistics for marketing and consumer research*. Sage.

Medina Moral, E. (2012). Modelos de elección discreta.

Meilán, J. J. G., Carro, J., Guerrero, C., Carpi, A., & Gómez, C. (2012). El efecto de memoria congruente con el estado afectivo: reconocimiento diferencial de palabras de tristeza y alegría. *Anales de psicología*, 28(1), 266.

Molinero L.(2001). Riesgo Relativo y Número Necesario a Tratar. *Asociación de la Sociedad Española de Hipertensión Liga Española para la lucha contra la hipertensión.*

Monge, J. F., & Pérez, A. J. (2000). Estadística no paramétrica: Prueba Chi-cuadrado X². *Barcelona: Universitat Oberta de Catalunya.*

Pich, J. (2000). The role of subvocalization in rehearsal and maintenance of rhythmic patterns. *The Spanish journal of psychology*, 3, 63-67.

Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1), 81-106.

Reyes N.(2015) *Música: papel de respuestas afectivas en el condicionamiento clásico de preferencias*. Universidad de Los Andes, Bogotá D.C.

Rokach, L., & Maimon, O. (2005). Decision trees. *Data mining and knowledge discovery handbook*, 165-192.

Scherer, K. R., & Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and emotion*, 1(4), 331-346.

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge university press.

Varela, J., & Rial, A. (2008). Estadística práctica para la investigación en ciencias de la salud.

Vázquez Villazón, M., & Bernard Cerezo, M. E. D. (1992). Métodos estadísticos aplicados a la investigación psicológica. *Revista Cubana de Psicología*, 9(2), 65-71.