

Instituto Tecnológico y de Estudios Superiores de Occidente

Reconocimiento de validez oficial de estudios de nivel superior según acuerdo secretarial 15018, publicado en el Diario Oficial de la Federación del 29 de noviembre de 1976.

Departamento de Electrónica, Sistemas e Informática
Maestría en Sistemas Computacionales



Analítica de preguntas con respuestas tipo Likert en encuestas de satisfacción de cursos de Algoritmos y Programación

TRABAJO RECEPCIONAL que para obtener el **GRADO** de
MAESTRO EN SISTEMAS COMPUTACIONALES

Presenta: **CAROLINA SANTANA DE LOS RÍOS**

Asesor: **GISEL CHAVEZ HERNANDEZ**

Tlaquepaque, Jalisco. Julio 2022.

TABLA DE CONTENIDO

1. INTRODUCCIÓN	6
1.1. ANTECEDENTES	6
1.2. JUSTIFICACIÓN	6
1.3. PROBLEMA	7
1.4. OBJETIVOS	7
1.4.1. OBJETIVO GENERAL	7
1.4.2. OBJETIVOS ESPECÍFICOS	7
1.5. NOVEDAD CIENTÍFICA, TECNOLÓGICA O APORTACIÓN	7
2. MARCO TEÓRICO	8
2.1. MÉTODOS DE MEDICIÓN EN PSICOLOGÍA	8
2.2. CONSTRUCTO	8
2.3. ESCALAS LIKERT	8
2.4. CORRELACIONES	9
2.4.1. COEFICIENTE DE CORRELACIÓN DE PEARSON	9
2.4.2. COEFICIENTE DE CORRELACIÓN DE SPEARMAN	10
2.4.3. COEFICIENTE DE CORRELACIÓN DE KENDALL	10
2.5. ANÁLISIS FACTORIAL	11
2.5.1. PRUEBA DE KAISER-MEYER-OLKIN	12
2.5.2. PRUEBA DE BARTLETT	12
2.5.3. MINRES	12
2.5.4. ROTACIONES	12
2.5.4.1. ROTACIÓN VARIMAX	13
2.5.4.2. ROTACIÓN QUARTIMAX	13
2.5.4.3. ROTACIÓN OBLIMIN	13
2.5.4.4. ROTACIÓN EQUAMAX	13
2.5.4.5. ROTACIÓN PROMAX	13
2.6. ALFA DE CRONBACH	14
2.7. ALFA ORDINAL	14
2.8. METODOLOGÍA CRISP-DM	14
3. ESTADO DEL ARTE	16
3.1. COEFICIENTES DE CONFIABILIDAD ALFA ORDINAL Y ALFA DE CRONBACH	16
3.2. ANÁLISIS FACTORIAL	16
3.3. ANÁLISIS CONFIRMATORIO	17
4. RESULTADOS Y DISCUSIÓN	19
4.1. RESULTADOS DE LA FASE DE COMPRESIÓN DEL NEGOCIO	19
4.2. RESULTADOS DE LA FASE DE COMPRESIÓN DE LOS DATOS	20
4.2.1. DIAGRAMAS DE DISPERSIÓN	23
4.2.2. HISTOGRAMAS	24
4.2.3. ESTUDIO EXPLICATIVO DEL COMPORTAMIENTO DE LOS CONSTRUCTOS EN EL TIEMPO	26
4.2.3.1. CONSTRUCCIÓN DE LAS SERIES DE TIEMPO	26
4.2.3.2. SERIES DE TIEMPO CON VALORES DE MEDIAS	26
4.2.3.3. GRÁFICAS DE CAJAS Y BIGOTES (BOXPLOT)	30
4.3. RESULTADOS DE LA FASE DE PREPARACIÓN DE LOS DATOS	34
4.3.1. MAPA DE CLÚSTER O DENDOGRAMA	38
4.4. RESULTADOS DE LA FASE DE MODELADO	41
4.4.1. ANÁLISIS FACTORIAL EXPLORATORIO	41
4.4.2. ANÁLISIS FACTORIAL CONFIRMATORIO (CFA)	45
4.4.2.1. PRUEBA DE FIABILIDAD CON EL ALFA DE CRONBACH	45

LISTA DE FIGURAS

Figura 1 - Histograma Constructo 1. Fuente: elaboración propia	25
Figura 2 - Histograma Constructo 2. Fuente: elaboración propia	25
Figura 3 - Histograma Constructo 3. Fuente: elaboración propia	25
Figura 4 - Serie de tiempo con valores de medias para el Constructo 1. Fuente: elaboración propia	26
Figura 5 - Tendencia, estacionalidad y ruido, con valores de medias para el Constructo 1. Fuente: elaboración propia	27
Figura 6 – Medias móviles para el Constructo 1.. Fuente: elaboración propia	27
Figura 7 - Serie de tiempo con valores de medias para el Constructo 2. Fuente: elaboración propia	28
Figura 8 - Tendencia, estacionalidad y ruido con valores de medias para el Constructo 2. Fuente: elaboración propia	28
Figura 9 – Medias móviles para el Constructo 2. Fuente: elaboración propia	28
Figura 10 - Serie de tiempo con valores de medias para el Constructo 3. Fuente: elaboración propia	29
Figura 11 - Tendencia, estacionalidad y ruido con valores de medias para el Constructo 3. Fuente: elaboración propia	29
Figura 12 – Medias móviles para el Constructo 3. Fuente: elaboración propia	30
Figura 13 - Boxplot del Constructo 1. Fuente: elaboración propia	31
Figura 14 - Boxplot del Constructo 2. Fuente: elaboración propia	32
Figura 15 - Boxplot del Constructo 3 Fuente: elaboración propia	33
Figura 16 - Mapa de calor del Constructo 1. Fuente: elaboración propia	34
Figura 17 - Mapa de calor del Constructo 2. Fuente: elaboración propia	36
Figura 18 - Mapa de calor del Constructo 3. Fuente: elaboración propia	37
Figura 19 - Clústeres del Constructo 1. Fuente: elaboración propia	39
Figura 20 - Clústeres del Constructo 2. Fuente: elaboración propia	40
Figura 21 - Clústeres del Constructo 3. Fuente: elaboración propia	41
Figura 22 - Mapa de clúster de EFA con rotación varimax. Fuente: elaboración propia	42
Figura 23 - Mapa de clúster de EFA con rotación promax. Fuente: elaboración propia	44
Figura 24 - Mapa de clúster del EFA sin rotación. Fuente: elaboración propia	45

LISTA DE TABLAS

Tabla 1 - Ejemplo de escalas Likert. Fuente [5].....	9
Tabla 2 - Componentes para considerar préstamos a estudiantes. Fuente: [11].....	17
Tabla 3 - Librerías de Python utilizadas. Fuente: elaboración propia	19
Tabla 4 - Campos de la fuente de datos. Fuente: elaboración propia	20
Tabla 5 - Preguntas cerradas. Fuente: elaboración propia.....	22
Tabla 6 - Encuestas, profesores y grupos en IAE 2016 – 2019. Fuente: elaboración propia	22
Tabla 7 - Grupos con diferente cantidad de profesores. Fuente: elaboración propia	23
Tabla 8 - Relación Respuesta(renglón) y Pregunta(columna). Fuente: elaboración propia	23
Tabla 9 - Tabla de frecuencia con totales. Fuente: elaboración propia	24
Tabla 10 - Tabla de frecuencia con totales. Fuente: elaboración propia	24
Tabla 11 - Series de tiempo con datos estadísticos. Fuente: elaboración propia.....	26
Tabla 12 - Serie del Constructo 1. Fuente: elaboración propia	31
Tabla 13 - Serie del Constructo 2. Fuente: elaboración propia	32
Tabla 14 - Series del Constructo 3. Fuente: elaboración propia.....	33
Tabla 15 - Correlaciones importantes del Constructo 1 “Satisfacción con el curso”. Fuente: elaboración propia	34
Tabla 16 - Correlaciones importantes del Constructo 2 “Satisfacción con el desempeño del profesor”. Fuente: elaboración propia	36
Tabla 17 - Correlaciones importantes del Constructo 3 “Satisfacción general con el curso”. Fuente: elaboración propia	37
Tabla 18 - Análisis factorial con diferentes rotaciones. Fuente: elaboración propia	42

1. INTRODUCCIÓN

Tener encuestas donde un estudiante puede evaluar el desempeño de un profesor sobre un curso en particular, es una herramienta útil para analizar si tanto el profesor como la clase y curso impartido fueron buenos desde el punto de vista del estudiante. Por esta razón, es importante contar con una herramienta que realmente permita obtener la opinión del estudiante, es decir, que permita obtener datos que sean de calidad para determinar si el profesor realizó un buen trabajo para impartir sus clases y el curso en general. El ITESO pide a sus estudiantes evaluar a sus profesores en diferentes momentos durante el curso: durante las primeras semanas y al final del periodo, de esta forma se obtiene la percepción del estudiante al inicio y fin del curso, todo con el fin de obtener datos de calidad que permitan identificar si el profesor realizó un buen trabajo. Es importante para el ITESO proporcionar educación de calidad para sus estudiantes y estas encuestas sirven como indicador para dicho propósito.

1.1. Antecedentes

El IAE tiene una sección de preguntas con respuesta tipo Likert, a partir de las cuales se persigue cuantificar los siguientes constructos (variables latentes): “satisfacción con las actividades del curso”, “satisfacción con el profesor” y “satisfacción general con el curso”. Dichas evaluaciones son aplicadas antes de finalizar los cursos del periodo escolar.

El IAE se comenzó a aplicar desde 2012, pero en su momento no se probó que con las preguntas diseñadas era válido inferir sobre los constructos antes mencionados, ni se hicieron estudios de analítica descriptiva ni de diagnóstico acerca de sus resultados. En este contexto, en 2019 se dio inicio a un proyecto de mejora en el IAE que incluyó la validación del mismo, así como el análisis de sus resultados a través de varios semestres. Es en este contexto que decidimos tomar como tema del TOG la analítica de respuestas tipo Likert del IAE para los cursos de Algoritmos y Programación.

1.2. Justificación

La estructura de preguntas y respuestas implementada actualmente por IAE solo permite sacar conclusiones por pregunta y además no todas son realmente útiles para conocer la opinión de los estudiantes sobre la satisfacción con las actividades del curso, con el profesor y con el curso en general. Con este trabajo se probará la validez de dicha encuesta para inferir sobre los constructos ya mencionados, así como un análisis del número adecuado de preguntas con respuestas tipo Likert que permitirán medir los constructos. Adicionalmente, la Coordinación docente de Algoritmos y Programación contará con un estudio descriptivo y de diagnóstico de lo ocurrido desde 2016 a 2019, lo cual ayudará en la toma de decisiones para la mejora de dicha materia.

1.3. Problema

Actualmente el IAE tiene encuestas con respuestas tipo Likert, sin embargo, no hay evidencia que afirme que la actual agrupación de preguntas permita medir los 3 constructos ya referidos. Las primeras analíticas hechas solo sacaron unas pocas conclusiones por pregunta utilizando la media muestral como estadístico. Otro problema es que alumnos y académicos consideran muy larga la encuesta y se presume que hay preguntas que son innecesarias.

1.4. Objetivos

1.4.1. Objetivo General

Determinar la fiabilidad de la actual IAE y genera una nueva propuesta de encuesta con el mínimo de preguntas que requiera cada constructo

1.4.2. Objetivos específicos

- Realizar un estudio exploratorio de los datos del IAE para la materia Algoritmos y Programación, haciendo una descripción de los resultados entre 2016 y 2019
- Evaluar el instrumento actual para tener un diagnóstico en cuanto a si mide o no los constructos que asegura evaluar
- Obtener un nuevo instrumento que demuestre validez en cuanto a la medición de los constructos y más pequeño.

1.5. Novedad científica, tecnológica o aportación

Con la nueva encuesta para IAE, ITESO tendrá una herramienta validada con técnicas de analítica de datos y pruebas de fiabilidad estadística que le permitirá realizar un mejor análisis sobre la opinión de sus estudiantes relacionada al curso y su profesor.

Este trabajo busca confirmar la validez del instrumento de evaluación de los cursos del ITESO, llamado Instrumento de Apreciación Estudiantil (IAE), el cual tiene una sección de respuestas tipo Likert, a partir de las cuales se persigue cuantificar los siguientes constructos: “satisfacción con las actividades del curso”, “satisfacción con el profesor” y “satisfacción general con el curso”. El proceso para confirmar la validez del instrumento incluye las medidas de adecuación muestral Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Bartlett, el análisis factorial exploratorio con diferentes rotaciones y el análisis factorial confirmatorio con el cálculo de alfa de Cronbach para confirmar la validez y consistencia interna de los tres constructos. El conjunto de datos utilizados para realizar el análisis son datos históricos de la materia “Algoritmos y programación” entre los periodos primavera 2016 a primavera 2019.

A continuación se presenta en el capítulo 2 un marco teórico que contiene las investigaciones, teorías y conceptos claves de nuestro trabajo.

2. MARCO TEÓRICO

Se expondrán las investigaciones, teorías y conceptos claves de nuestro trabajo, o sea, las relativas a probar la validez de una encuesta con preguntas de respuesta tipo Likert a partir de la cual se cuantifican (se miden indirectamente) constructos.

2.1. Métodos de medición en Psicología

En 1946, el psicólogo Stanley Smith Stevens desarrolló una teoría de escalado y medición y, proponiendo cuatro escalas que actualmente se siguen utilizando: nominal, ordinal, intervalo y proporciones. El proceso para asignar una escala consiste en relacionar un número a objetos con base a una regla [1], gracias a esta asignación de escalas es posible realizar la medición de un fenómeno.

Se han desarrollado diferentes métodos de medición, los más comunes son: escalas *Thurstone*, escalas *Guttman* y escalas *Likert* [2]. Cada una tiene sus propios procesos, así como pros y contras. Sin embargo, las escalas Likert son el método de medición más común en contextos educativos [3]. Es por ello que el IAE del ITESO empleó dicha escala y, por tanto nuestro trabajo las tiene como objeto de estudio en su correcta aplicación e interpretación.

2.2. Constructo

El profesor y psicólogo británico John F. Binning, define a los constructos como una herramienta para facilitar la comprensión del comportamiento humano y resolver un problema científico. Su nombre deriva de una construcción mental, pero un constructo va más allá de la ideación y el proceso social que implica la comunicación. En matemáticas, se considera a los constructos como objetos autónomos, incluso cuando éstos no existen físicamente [4]. En este trabajo se analiza la validez de la cuantificación de 3 constructos (variables latentes), a partir de preguntas con respuestas tipo Likert.

2.3. Escalas Likert

Las escalas Likert proponen que un fenómeno (constructo) puede medirse mediante una clasificación de los sentimientos, aptitudes, ideas o percepciones de un individuo y después relacionarla a un conjunto de ítems (que pueden ser preguntas de una encuesta) [5]. La Tabla 1 muestra preguntas que implementan la escala Likert misma que consiste en un conjunto de ítems utilizados para medir una característica o constructo (en este caso la satisfacción general con un

profesor), donde cada uno de éstos cuenta con una colección de respuestas acompañadas de números que entre sí tienen una distancia equivalente.

Tabla 1 - Ejemplo de escalas Likert.
Fuente [5]

Escala Likert

	Muy en desacuerdo	Desacuerdo	Neutral	De Acuerdo	Muy de Acuerdo
El profesor nos inspira para trabajar duro	1	2	3	4	5
El profesor proporciona la ayuda necesaria	1	2	3	4	5
El profesor permite a los demás expresar sus ideas	1	2	3	4	5

Es importante resaltar que un solo ítem que utiliza el formato de respuesta Likert, no debe considerarse como escala Likert ya que ésta consiste en un conjunto de elementos [5]. También es importante diferenciar las “escalas tipo Likert”; John S. Uebersax, psicólogo y bioestadístico, llamó así al incorrecto uso de escalas Likert como medida de frecuencia, es decir, un conjunto de ítems cuyos valores son sumados, presentados horizontalmente, espaciados con números enteros y con etiquetas igualmente espaciadas. Se dice que es incorrecto porque las escalas Likert deben ser utilizadas como un estricto método de medición al nivel de aprobación por parte de un individuo [6].

2.4. Correlaciones

El coeficiente de correlación depende de la covarianza, su valor es igual a la covarianza dividida entre el producto de las desviaciones estándar de las variables. Por lo tanto, una covarianza positiva siempre producirá una correlación positiva y una covarianza negativa producirá una correlación negativa [7].

2.4.1. Coeficiente de correlación de Pearson

Un coeficiente de correlación mide el grado en que dos variables tienden a cambiar, describe la fuerza y dirección de la relación. El coeficiente de correlación de Pearson permite medir la fuerza o grado de asociación entre dos variables aleatorias cuantitativas que poseen una distribución normal conjunta y se define por la fórmula (1) [7].

$$\rho = \frac{cov(x,y)}{\sigma_x \sigma_y} \quad (1)$$

Donde:

cov = covarianza poblacional

$\sigma_x \sigma_y$ = desviaciones estándar de la población

Los siguientes supuestos fundamentan el coeficiente de correlación de Pearson [7]:

1. La distribución conjunta de las variables (X, Y) debe ser normal
2. Observar que cada variable se distribuya en forma normal, si una variable se desvía de la normalidad, la distribución conjunta tampoco es normal
3. Existe una relación tipo lineal entre las variables (X, Y)
4. Para cada valor de X, hay una subpoblación de valores Y normalmente distribuidas
5. Subpoblaciones de valores Y tienen varianza constante
6. Las medias de las subpoblaciones de Y se encuentran en la misma línea recta
7. Subpoblaciones de X tienen varianza constante
8. Las medias de subpoblaciones de X se encuentran en la misma línea recta

Para cada valor de Y hay una subpoblación de valores X que están normalmente distribuidas

2.4.2. Coeficiente de correlación de Spearman

Coeficiente no paramétrico alternativo al coeficiente de correlación de Pearson cuando éste cumple los siguientes supuestos [7]:

1. La distribución conjunta de las variables (X, Y) debe ser normal
2. Observar que cada variable se distribuya en forma normal, si una variable se desvía de la normalidad, la distribución conjunta tampoco es normal
3. Subpoblaciones de valores Y tienen varianza constante
4. Las medias de las subpoblaciones de Y se encuentran en la misma línea recta
5. Subpoblaciones de X tienen varianza constante
6. Las medias de subpoblaciones de X se encuentran en la misma línea recta

Además, la correlación Spearman suele usarse para evaluar relaciones en donde hay variables ordinales, este coeficiente se calcula utilizando la fórmula (2) [7].

$$\gamma_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n} \quad (2)$$

Donde:

d_i = diferencia entre los correspondientes estadísticos de orden x – y

n = número de parejas de datos

2.4.3. Coeficiente de correlación de Kendall

Coeficiente utilizado cuando se estudia la relación entre variables cualitativas de tipo ordinal. Además, su aplicación tiene sentido si las variables objeto de estudio no poseen una distribución poblacional conjunta normal; es decir, si se requiere determinar el grado de asociación lineal entre las variables cuantitativas, pero no las mismas que siguen un comportamiento normal. Este indicador está basado en rangos y no en los datos originales, su estimación requiere que los valores

de la variable ordinal sean transformados en rangos, este coeficiente se ve poco afectado ante la presencia de un número pequeño de valores atípicos en la muestra estudiada [8].

Los resultados reportados por el coeficiente de Kendall son más bajos con respecto a los coeficientes Spearman y Pearson, en aquellas situaciones donde se analizan las asociaciones lineales con la misma intensidad (sin valores atípicos). No quiere decir que los resultados de Kendall son menos precisos, esto es debido a que trabaja con valores basados en rangos. Los resultados también reportan la misma tendencia en datos con distribución monótonas no lineales, en comparación con Spearman [8]. La fórmula (3) muestra la fórmula para el cálculo de la correlación de Kendall sin observaciones empatadas [8].

$$T_{XY} = \frac{2 * S}{n (n - 1)} \quad (3)$$

Donde:

S = P – M

P = Número de valores positivos o “acuerdos”

M = Número de valores negativos o “desacuerdos”

n = Tamaño de la muestra

2.5. Análisis factorial

El análisis factorial es el método de elección para interpretar los cuestionarios de auto informe y es utilizado comúnmente en los campos de la psicología y educación. El análisis factorial es un procedimiento estadístico multivariado que tiene diversos usos, reduce un gran número de variables en pequeños conjuntos de variables o factores, establece dimensiones subyacentes entre variables medidas y constructos latentes, y proporciona evidencia sobre la validez de constructo con escalas de auto informe [9].

Existen dos tipos principales de análisis factorial: análisis factorial exploratorio (*Exploratory Factor Analyzer, EFA*) y análisis factorial confirmatorio (*Confirmatory Factor Analyzer, CFA*). Al trabajar con EFA, el investigador no tiene expectativas sobre el número de variables, EFA le permite explorar las principales dimensiones para generar una teoría o un modelo a partir de un conjunto de constructos latentes comúnmente representados por un conjunto de ítems. Por otro lado, con CFA el investigador prueba una teoría propuesta, tiene suposiciones y expectativas basadas en la teoría a priori con respecto al número de factores y qué teorías o modelos de factores se ajustan mejor [9].

Al trabajar con EFA, se sigue un protocolo de 5 pasos que proporciona a los investigadores un punto de referencia inicial. El primer paso es identificar si los datos son apropiados para el análisis de factores, para ello se puede calcular Kaiser-Meyer-Olkin (KMO) y la prueba de Bartlett. El segundo paso es identificar cómo serán extraídos los factores, por ejemplo, análisis de componentes principales (PCA), MINRES, entre otros. El tercer paso es identificar el criterio para determinar la extracción de factores, se pueden utilizar las propiedades de valores propios de PCA (*Eigenvalues Decomposition, EVD*) y posteriormente se puede calcular valor singular de descomposición (*Singular Value Decomposition, SVD*). El cuarto paso es seleccionar el método de rotación para maximizar las altas cargas de los ítems y minimizar las cargas bajas, además de

producir una solución más interpretable y simplificada. El quinto paso es la interpretación, el investigador examina cuáles variables son atribuibles a un factor y se le asigna un nombre. Al menos dos o tres variables deben cargar un factor para que pueda tener una interpretación significativa [9].

Después de EFA, es necesario confirmar los resultados obtenidos a través de CFA, la cual es una técnica que busca confirmar que el número de factores o constructos y las cargas observadas se ajustan a lo esperado en la teoría. Por lo tanto, es necesario evaluar la fiabilidad y validez de la escala. El método más común es el cálculo del alfa de Cronbach, el cual mide la fiabilidad y la consistencia interna [10].

2.5.1. Prueba de Kaiser-Meyer-Olkin

La prueba de Kaiser-Meyer-Olkin (KMO) mide cuán adecuados son los datos para el análisis de factores utilizando la fórmula (4). Como resultado se obtiene un valor en el rango de 0 a 1, donde un valor mayor a 0.5 es considerado lo mínimo apropiado para el análisis factorial [11].

$$KMO = \frac{\sum r^2_{ij}}{\sum r^2_{ij} + \sum u} \quad (4)$$

Donde:

$R = [r_{ij}]$ es la matriz de correlación

$U = [u_{ij}]$ es la matriz de covariancia parcial

2.5.2. Prueba de Bartlett

La prueba de Bartlett es una medida de cuán adecuados son los datos para el análisis de factores, compara la matriz de correlación observada con la matriz de identidad, será significativa para el análisis de factores si el valor resultante $p < 0.05$ [7], el cual indica que la matriz de correlación no es una matriz de identidad, por lo tanto, los valores son adecuados para el análisis factorial [12].

2.5.3. MINRES

MINimum RESiduals (MINRES) fue introducido por Harry Harman (1960) como una forma de realizar el análisis factorial por mínimos cuadrados en donde solo se requiere ajustar una matriz de cargas factoriales a la matriz de correlación [13].

2.5.4. Rotaciones

Las rotaciones minimizan la complejidad de las cargas factoriales simplificar la interpretación de la estructura resultante. Las matrices de carga factorial no son únicas, para cualquier solución que involucre dos o más factores existe un número infinito de orientaciones de los factores que explican igual de bien los datos originales. La rotación de las matrices de carga intenta dar una solución con la mejor y más simple estructura. Existen dos tipos de rotación que son ortogonal y oblicua [14].

Las rotaciones ortogonales restringen los factores para que no estén correlacionados. En muchos casos no es realista esperar que los factores no estén correlacionados y obligarlos a no estar

correlacionados hace que sea menos probable que la rotación produzca una solución con una estructura simple. Por otro lado, las rotaciones oblicuas permiten que los factores se correlacionen entre sí, por lo general producen soluciones con una estructura más simple [14]. Dependiendo del entendimiento que se tenga sobre los ítems se puede seleccionar un tipo de rotación, por ejemplo, si se considera que los ítems representan múltiples variables latentes y éstas no pueden estar correlacionadas, entonces es recomendado utilizar la rotación varimax [15].

2.5.4.1. Rotación varimax

La rotación varimax es una técnica estadística utilizada en un nivel de análisis factorial como un intento para aclarar la relación entre los factores. Es una rotación ortogonal que tiene por objetivo maximizar la varianza compartida entre los elementos. Para maximizar la varianza se debe incrementar la correlación al cuadrado de los elementos relacionados a un factor, mientras se disminuye la correlación de cualquier otro factor. Básicamente, la rotación varimax simplifica las cargas de elementos removiendo el término medio e identificando específicamente al factor sobre el cual se cargan los datos [16].

2.5.4.2. Rotación quartimax

Método de rotación ortogonal que minimiza el número de factores necesarios para explicar cada variable. Este método simplifica la interpretación de las variables observadas [17].

2.5.4.3. Rotación oblimin

Rotación oblicua que permite y calcula las correlaciones entre los factores latentes, produce un resultado más complejo que el proporcionado por la rotación varimax. Además del factor de correlación, el factor de la matriz de cargas (la relación entre los factores latentes y los indicadores observados) se descompone en la matriz de patrón y estructura factorial que representan los efectos directos y totales de los factores latentes [15].

2.5.4.4. Rotación equamax

Método de rotación ortogonal resultante al combinar el método varimax, el cual simplifica los factores, y el método quartimax, el cual simplifica las variables. Se minimiza el número de variables que cargan mucho en un factor y el número de factores necesarios para explicar una variable [17].

2.5.4.5. Rotación promax

Método de rotación oblicua que tiene la característica de ser la más rápida y conceptualmente simple de las rotaciones oblicuas. Realiza dos pasos, en el primero define la matriz objetivo, muchas veces obtenida del resultado de la rotación varimax; el segundo paso es calcular un ajuste de mínimos cuadrados de la solución varimax a la matriz objetivo. Los resultados obtenidos pueden ser interpretados observando las correlaciones entre los ejes de rotación y las variables originales, dichas correlaciones son interpretadas como cargas [17].

2.6. Alfa de Cronbach

Uno de los objetivos de este trabajo es determinar la fiabilidad de los tres constructos ya mencionados, se escogió el alfa de Cronbach por las bondades que tiene dicho coeficiente para medir la validez. El coeficiente alfa es la medida más popular para la confiabilidad interna, introducida por Guttman (1945) [18] y popularizada por Cronbach (1951) [19]. Este asume que los ítems tienen igual importancia cuando son medidos en el constructo no observado, pero se permite que las variaciones de error respectivas para cada elemento difieran [20].

Lee J. Cronbach denominó Alfa de Cronbach a un coeficiente para medir la fiabilidad de una escala de medida. Básicamente, se trata de una media de las correlaciones entre variables que forman parte de una escala, puede ser expresado utilizando una notación matricial mostrada en la fórmula (5) [19]

$$\alpha = \frac{n}{n-1} \left(1 - \frac{tr(\Sigma)}{1'\Sigma 1} \right) \quad (5)$$

Donde:

Σ = matriz de varianza-covarianza de $n \times n$ entre los elementos de la prueba

$tr(\Sigma)$ = traza de la matriz

Cuando Σ es reemplazado por la matriz de correlación R , o alternativamente cada ítem de prueba es primero estandarizado para tener una media de cero y varianza de uno antes de calcular $\hat{\Sigma}$, entonces se obtendrá la estimación del alfa estandarizado. El alfa estandarizado asume que la varianza de cada ítem es igual, resultando en el más estricto y menos realista, supuesto que los ítems son paralelos [20]

2.7. Alfa ordinal

Valor conceptualmente equivalente al alfa de Cronbach. El alfa ordinal se basa en la matriz de correlación policórica, logrando una mejor estimación de alfa mediante mediciones con datos ordinales, se calcula utilizando la fórmula (6), aplicándola sobre la matriz de correlaciones policóricas [21]

$$\alpha = \frac{n}{n-1} \left[\frac{n(\bar{\lambda})^2 - \overline{\lambda^2}}{n(\bar{\lambda})^2 - \overline{u^2}} \right] \quad (6)$$

Donde:

n = número de ítems

$\bar{\lambda}$ = media aritmética de los pesos factoriales

$\overline{\lambda^2}$ = media aritmética de los cuadrados de los pesos factoriales

$\overline{u^2}$ = media aritmética de las unicidades de los pesos factoriales

2.8. Metodología CRISP-DM

Cross-Industry Process for Data Mining (CRISP-DM), creado por la iniciativa European Strategic Program on Research in Information Technology (ESPRIT), es un modelo de procesamiento con

las siguientes seis fases que permiten planificar, organizar e implementar un proyecto de ciencia de datos [22].

Fase 1. Inicio y planeación del Proyecto

1. Definir salidas del Proyecto
2. Evaluar la situación actual
3. Determinar las metas
4. Producir el plan del Proyecto

Fase 2. Comprensión de los datos

1. Describir los datos
2. Explorar los datos
3. Verificar la calidad de los datos
4. Reportar la calidad de los datos
5. Integrar resultados en un reporte de recolección de datos iniciales

Fase 3. Preparación de los datos

1. Seleccionar datos
2. Limpiar datos
3. Construir datos requeridos
4. Integrar datos

Fase 4. Modelado – análisis factorial

1. Seleccionar técnica de modelado
2. Generar diseño de prueba
3. Construir el modelo
4. Valorar el modelo

Fase 5. Evaluación

1. Evaluar resultados
2. Revisar el proceso
3. Determiner próximos pasos

Fase 6. Despliegue (implementación)

1. Planear despliegue
2. Planear monitoreo y mantenimiento
3. Producir reporte final

Las fases 1-5 se reportan en este trabajo de obtención de grado

3. ESTADO DEL ARTE

A continuación, se presenta el estado del arte y se incluyen artículos relacionados a encuestas que implementaron escalas tipo Likert, cálculo de confiabilidad usando el alfa Cronbach y el análisis de factores.

3.1. Coeficientes de confiabilidad alfa ordinal y alfa de Cronbach

Zumbo propuso un nuevo estadístico de confiabilidad, el cual llamó “alpha ordinal”, creado para dar cuenta de la naturaleza categórica de la respuesta del ítem que se encuentra comúnmente en las pruebas educativas, encuestas psicológicas, instrumentos de medición clínicos, escalas de medición, entre otros. Sugiere reemplazar la correlación Pearson o Spearman con una correlación policorica estimada o tetracorica. La justificación para esta sustitución es que la estimación de confiabilidad tradicional, subestima la verdadera confiabilidad de la población porque correlaciones observadas entre variables categóricas suelen ser más bajas que las correlaciones entre variables latentes continuas subyacentes a partir de las cuales se manifestaron las variables categóricas. Además, argumenta que las correlaciones Pearson y Spearman requieren datos continuos y las estimaciones de correlación deben reemplazarse por estimaciones que no se basen únicamente en relaciones de datos categóricos. Concluyó que el alpha ordinal es más óptimo que el coeficiente del alpha de Cronbach cuando los datos de respuesta son categóricos, y recomiendan utilizar el alpha ordinal en lugar del coeficiente del alpha de Cronbach en la práctica donde sea que datos categóricos son recolectados [21].

Sin embargo, el coeficiente del alfa de Cronbach nunca ha requerido datos continuos a nivel de ítem, solo requiere que las relaciones observadas entre cada ítem sean codificadas en formato de intervalos [23], los cuales no necesitan ser codificadas con decimales o fracciones, solo se requiere que las distancias entre los valores representen la misma cantidad. Esta propiedad ha generado confusión sobre los requerimientos de datos para el coeficiente del alfa de Cronbach debido a que la codificación de las variables ordinales no incluye valores decimales o fraccionarios, y Zumbo concluye erróneamente que las estimaciones de covarianza de Pearson o Spearman no son válidas [24].

3.2. Análisis factorial

Se realizó un proceso para revisar los criterios de consideración de préstamos estudiantiles mediante el método de análisis factorial, utilizando cinco pasos para determinar los factores que influyen en dicha consideración: Kaiser-Meyer-Olkin, prueba de esfericidad de Bartlett, Análisis de Componentes Principales (PCA), rotación (varimax), evaluación e interpretación. Para esta investigación se utilizó un conjunto de datos históricos de los candidatos a préstamos estudiantiles

entre 2014 y 2015, con un total de 1146 muestras con siete criterios, se tuvo lo necesario para que el análisis factorial pudiese reducir las dimensiones de la varianza en la muestra al identificar los factores significativos para la consideración de préstamos estudiantiles [11].

La Tabla 2 muestra los componentes considerados para esta investigación. Se obtuvo un KMO > 0.5, considerado como aceptable para el análisis factorial. En la prueba de Bartlett un valor significativo $p < 0.5$, lo que confirma que se puede aplicar el análisis factorial sobre el conjunto de datos. Se realizó el Análisis de Componentes Principales (PCA) porque es el método por defecto para la extracción de muchos paquetes de software como SPSS. Posteriormente, se aplicó la rotación varimax porque minimiza el número de variables que tienen cargas altas en cada factor [11].

Tabla 2 - Componentes para considerar préstamos a estudiantes.
Fuente: [11]

Attribute code	Attribute description	Variable meaning
X ₁	Supporter who provide a financial for student	1 = father and mother, 2 = only father or mother, 3 = relative, parent 4 = person 5 = no Supporter
X ₂	Family's income	Income = income/1000
X ₃	Debt	0 = null, 1 = from occupation 2 = from education 3 = from occupation and education
X ₄	Education level of the sibling	0 = null 1 = Pre-elementary and less than, 2 = high school, 3 = diploma 4 = bachelor and higher
X ₅	Amount of money the student got from other scholarship	scholarship = scholarship/1000
X ₆	Positive score from the committee	0-10
X ₇	Negative score from the committee.	-10-0

El análisis factorial reveló que solo había tres factores que influyen en la consideración del préstamo estudiantil. El factor en el componente 1 consiste en las variables “ingreso familiar”, “cantidad de dinero que el estudiante recibió de otra beca” y “nivel de educación del hermano”. Por otro lado, los factores en el componente 2 y el componente 3 solo tienen una pequeña influencia en la consideración del préstamo estudiantil basado en los valores propios (*eigenvalues*). Por lo tanto, el estudio muestra el hallazgo de que los factores dominantes para considerar el préstamo estudiantil se encuentran en el componente 1, esto es porque los *eigenvalues* encontrados fueron en el componente 1 fueron mayores a los encontrados en los componentes 2 y 3 [11].

3.3. Análisis Confirmatorio

Este estudio examinó los efectos de la satisfacción laboral y el equilibrio entre la vida laboral y personal en el desempeño de las tareas del personal académico en una institución pública de

educación superior seleccionada en Malasia. Se realizaron entrevistas con personal de la empresa para obtener información sobre los factores que afectan el desempeño de las tareas del personal académico [25].

Se generó una encuesta de 16 ítems para medir la satisfacción laboral en el lugar de trabajo, como recompensas y beneficios, reconocimientos, oportunidades de promoción, entre otros. Los participantes indicaron su respuesta a través de una escala tipo Likert de cinco puntos, donde 1 indica “totalmente en desacuerdo” y 5 indica “totalmente de acuerdo”. Se obtuvo un coeficiente de confiabilidad con el alfa de Cronbach de 0.85 para la satisfacción laboral. Por otro lado, se generó una encuesta de 7 ítems para medir el equilibrio de la vida laboral, utilizando una escala tipo Likert de 5 puntos, donde 1 es “muy poco” y 5 “mucho”. El alfa de Cronbach obtenido fue de 0.6 para el equilibrio de la vida laboral [25].

El hallazgo de la investigación mostró una asociación positiva entre la satisfacción laboral y el equilibrio entre la vida laboral y personal en el desempeño de las tareas del personal académico. Además, se obtuvo que la satisfacción laboral tiende a ser el predictor más fuerte en comparación con el equilibrio entre la vida laboral y personal [25].

Aplicación del análisis confirmatorio en un estudio de encuesta de satisfacción

Para determinar la confiabilidad y consistencia entre los ítems de las preguntas relacionadas a la evaluación de satisfacción de pasajeros y calidad del servicio de la terminal 2 del aeropuerto Murtala Muhammed, se realizó un estudio que explica la implementación de escalas tipo Likert y el análisis de alfa de Cronbach. La encuesta aplicada explora solo datos de confiabilidad y consistencia interna de la satisfacción de los pasajeros y calidad de servicio en el mercado internacional y la terminal 2 del aeropuerto Murtala Muhammed, Lagos, Nigeria.

Se realizó el análisis con el alfa de Cronbach para comprobar la fiabilidad y consistencia interna de las escalas tipo Likert, las cuales son adecuadas para estudios sociales y del comportamiento relacionadas a las percepciones, aptitudes, emociones, opiniones, personalidades y descripción del entorno de las personas. El análisis se realizó sobre 39 preguntas para la satisfacción de los pasajeros y 39 preguntas para la satisfacción sobre la calidad del servicio. Se determinó el tamaño de la muestra con un 92% de nivel de confianza y con un error máximo de 0.08 debido a la conciencia de los pasajeros aéreos, obteniendo como resultado una muestra de 120 encuestas. Se distribuyó dicha cantidad de encuestas y se obtuvo respuesta de 114, esto implica que se obtuvo una tasa de respuesta del 95%, excelente para el análisis de datos. Se realizó la prueba de fiabilidad para las declaraciones de 39 servicios de aeropuertos combinados con cinco atributos para medir la calidad de servicio y la satisfacción de los pasajeros, obteniendo un valor de 0.893 y 0.861 respectivamente. Estos resultados indican una alta confiabilidad del instrumento del cuestionario y consistencias internas de las escalas tipo Likert [26].

4. RESULTADOS Y DISCUSIÓN

A continuación, se presentan los resultados obtenidos de cada fase de la metodología CRISP-DM, el análisis de los datos, diferentes tipos de gráficas con sus respectivas interpretaciones, análisis factorial exploratorio con diferentes rotaciones y análisis factorial confirmatorio con el cálculo del alfa de Cronbach.

4.1. Resultados de la fase de comprensión del negocio

El lenguaje utilizado para realizar la exploración de datos es Python, la Tabla 3 muestra las librerías que fueron de utilidad para este trabajo. El código en Python fue ejecutado con ayuda de Jupyter Notebook [27] y Google Colab [28].

Tabla 3 - Librerías de Python utilizadas.
Fuente: elaboración propia

Herramienta	Descripción
factor_analyzer.factor_analyzer	Módulo para realizar exploración y análisis de factores con diferentes opciones de rotación. También incluye una clase para realizar un análisis de factores confirmatorio [29].
numpy	Paquete para el cómputo científico, provee arreglos multidimensionales, operaciones sobre los mismos como manipulación de forma, ordenamiento, selección, transformaciones de Fourier, entre otros [30].
pandas	Paquete que provee un alto rendimiento, una forma fácil de usar estructuras de datos y herramientas para el análisis de datos [31].
plotly	Librería para crear gráficas y visualizaciones interactivas. Incluye más de 40 tipos de gráficos de diferentes áreas como estadística, finanzas, geografía, ciencias, entre otros [32].
pingouin	Paquete estadístico con una lista de funciones estadísticas, por ejemplo, regresiones lineales, factores de bayes, fiabilidad y consistencia, pruebas de chi-cuadrada, entre otros [33].
scikit	Librería para el análisis de datos predictivo, permite realizar operaciones de clasificación, regresión, pre procesamiento, selección de modelos, entre otros [34].

sklearn. decomposition	Módulo que incluye algoritmos de descomposición de matrices. Incluye la clase de FactorAnalysis para realizar la exploración y análisis de factores con diferentes parámetros [29].
matplotlib	Librería para crear visualizaciones estáticas, animadas e interactivas en Python [35].
seaborn	Librería para la visualización de datos, con una interfaz de alto nivel para crear gráficos estadísticos, configura atributos visuales como color, tamaño y estilos, realiza transformaciones estadísticas para finalmente mostrar la gráfica con ejes, etiquetas y leyendas [36]
stats_model	Módulo de Python que provee clases y funciones para la estimación de modelos estadísticos, pruebas estadísticas y exploración estadística de datos [37]

Los algoritmos clave para la investigación son la prueba de KMO y la prueba de Bartlett, estas pruebas son de utilidad para evaluar si el conjunto de datos es adecuado para realizar el análisis factorial y se pueden calcular con la librería `factor_analyzer`. El EFA y el CFA son algoritmos clave para determinar si los tres constructos que se tienen para IAE (satisfacción con el profesor, satisfacción con el curso, satisfacción en general) son adecuados, pueden calcularse con la librería `factor_analyzer`. Para extraer los factores, se calculó MINRES mediante `factor_analyzer`. Para producir una solución más interpretable y simplificada sobre la carga de los ítems, se utilizaron rotaciones ortogonales y oblicuas, calculadas mediante `factor_analyzer`. Finalmente, el coeficiente del alfa de Cronbach para analizar la confiabilidad de los tres constructos del IAE, calculado con la librería `pingouin`.

4.2. Resultados de la fase de comprensión de los datos

La fuente de datos son archivos con extensión csv que incluyen los resultados de IAE entre 2016 – 2019. Cada año cuenta con dos periodos: “Primavera” y “Otoño”. Los archivos fueron anonimizados para guardar la confidencialidad de profesores y alumnos, la Tabla 4 incluye el listado de columnas incluidas en cada archivo csv, también indica cuáles fueron los campos que se utilizaron para los análisis posteriores.

Tabla 4 - Campos de la fuente de datos.
Fuente: elaboración propia

Nombre campo	Descripción
AREA_GENERAL	Valor por defecto: SP – Saberes profesionales
CALI_ALUM	Calificación obtenida por el alumno en la materia al concluir la semana 16
CODIGO	Identificador único del grupo en un periodo
COORDINADOR_DOC	Nombre del coordinador docente (valor anonimizado)

EVALUACION_ID	Identificador único de la encuesta por alumno – curso *Este campo se utilizó para los análisis
EVALUADO	Nombre del profesor (valor anonimizado)
FACTORAJUSTE	Valor estadístico de la universidad obtenido a partir de datos de años atrás y todos los cursos del ITESO
GRUPO_ID	Identificador único del curso independientemente del periodo *Este campo se utilizó para los análisis
PREGUNTA	Número de la pregunta *Este campo se utilizó para los análisis
PROGRAMAEDUCATIVO	Programa educativo al que pertenece el alumno
RESPUESTA	Es la apreciación que hace el alumno en la semana 13. Para las preguntas 2 – 17 se tienen las siguientes opciones: 1. Totalmente en desacuerdo 2. Parcialmente en desacuerdo 3. Indeciso 4. Parcialmente de acuerdo 5. Totalmente de acuerdo Para las preguntas 1 y 19 son respuestas abiertas. Este campo almacena datos en formato string
RESULTADO_PROFESOR	Calificación entre 0 y 100 interpolando en una línea recta que se genera usando datos de otros años
TIPO_EMPLEADO	Tipos de empleado: • PA – Profesor de Asignatura • TF – Profesor de tiempo fijo
TIPOPREGUNTA	Tipo de pregunta: • A – Abierta • C – Cerrada
VALORPARAPROMEDIO	Valores del 0 al 11: • 0 para preguntas abiertas (1 y 19) • 1 – 5 para la escala Likert • 6 en adelante para la pregunta 18 con las opciones *Este campo se utilizó para los análisis
PERIODO	Identificador del semestre *Este campo se utilizó para los análisis

Se tiene un total de 19 preguntas, sin embargo, no todas serán analizadas en este proyecto. La Tabla 5 incluye el listado de preguntas con opciones tipo Likert que serán consideradas para realizar los análisis necesarios.

Tabla 5 - Preguntas cerradas.
Fuente: elaboración propia

Pregunta	Contenido de la pregunta
2	Me han ayudado a identificar los posibles usos o aplicaciones que puede tener lo que aprendo
3	Han requerido que colabore con otros compañeros
4	Han promovido que aprenda a partir de mis propias acciones
5	Me han ayudado a reflexionar acerca de los aprendizajes que he ido logrando
6	Han requerido que aprenda mediante el uso de tecnologías de información y comunicación
7	Me han ayudado a relacionar con el contexto actual lo que aprendo en clase
8	Han requerido mi participación en situaciones que me permiten aprender a respetar distintas formas de ser y de pensar
9	Han requerido que consulte información en diversas fuentes
10	Me ha ayudado a tener un panorama claro acerca de lo que aprendería
11	Me ha ayudado a entender con claridad los criterios con los que sería evaluada(o)
12	Me ha ayudado a saber cuáles son los aspectos que debo mejorar
13	Me ha orientado para mejorar mi capacidad de expresar mis ideas de manera oral y/o escrita
14	Ha tomado en cuenta la opinión del grupo para mejorar sus clases
15	Ha mostrado dominio de los conocimientos y habilidades propios de esta asignatura
16	Recomiendo plenamente a este profesor
17	Cuál es tu nivel de satisfacción con los aprendizajes que has obtenido hasta el momento en esta asignatura

La Tabla 6 muestra algunos datos generales sobre periodos, total de encuestas, profesores y grupos que se tuvieron desde el 2016 hasta el 2019. En total se tuvieron 1187 encuestas, 20 profesores y 68 grupos. En la tabla anterior se puede observar que en otoño 2018 fue el periodo con más encuestas contestadas con 219, mientras que primavera 2016 fue el que menos tuvo con 132. Además, en otoño 2018 se tuvieron 11 profesores mientras que en primavera 2016 se tuvieron 7. En cuanto a grupos, los otoños de 2017 y 2018 tuvieron 12 grupos, siendo éstas la mayor cantidad, mientras que las primaveras 2018 y 2019 tuvieron 10 grupos. Los otoños presentan el mayor número de encuestas, profesores y grupos en comparación con las primaveras.

Tabla 6 - Encuestas, profesores y grupos en IAE 2016 – 2019.
Fuente: elaboración propia

Periodo	Encuestas	Profesores	Grupos
---------	-----------	------------	--------

O2016	132	7	7
O2017	200	8	12
O2018	219	11	12
P2016	118	9	8
P2017	199	9	9
P2018	147	8	10
P2019	172	10	10

La Tabla 7 muestra los grupos de tal forma que se puede observar la cantidad de grupos con más de un profesor, todo esto organizado por periodo. Se puede observar que otoño 2016, primavera 2016 y primavera 2016 tuvieron un grupo con dos profesores.

Tabla 7 - Grupos con diferente cantidad de profesores.
Fuente: elaboración propia

Periodo	Grupos	Profesores
O2016	6	1
O2016	1	2
O2017	12	1
O2018	12	1
P2016	1	2
P2016	7	1
P2017	1	2
P2017	8	1
P2018	10	1
P2019	10	1

4.2.1. Diagramas de dispersión

Utilizando los campos PREGUNTA y VALORPARAPROMEDIO, se generaron tablas cuyas columnas representan a las preguntas de cada constructo y los renglones representan a cada estudiante que contestó la encuesta. La Tabla 8 muestra un ejemplo de los datos de las preguntas 2 - 17.

Tabla 8 - Relación Respuesta(renglón) y Pregunta(columna).
Fuente: elaboración propia

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
0	4	5	5	3	5	4	4	5	3	4	4	4	2	2	1	2
1	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
2	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
3	2	2	2	2	2	2	2	3	2	1	2	3	4	4	3	3
4	4	4	3	4	4	3	3	3	3	3	3	3	3	3	3	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Para observar la información desde otro ángulo, se creó la Tabla 9 de frecuencias y la Tabla 10 de porcentajes, ambas tienen como factores las columnas anteriormente mencionadas. Nótese que las respuestas más utilizadas fueron las opciones 4 “Parcialmente de acuerdo” y 5 “Totalmente de acuerdo”, con un valor mayor a 100 en todas las preguntas. Los valores más altos para la opción 3 “Indeciso” se encuentran en las preguntas 8, 9, 12, 13, 17; los dos primeros valores corresponden al constructo 1 “satisfacción con el curso” y los demás al constructo 2 “satisfacción con el desempeño del profesor”. Por lo tanto, para dichas preguntas, hubo varios alumnos que estuvieron indecisos. Los valores para las opciones 1 “Totalmente en desacuerdo” y 2 “Parcialmente en desacuerdo” fueron bajos en todas las preguntas.

Tabla 9 - Tabla de frecuencia con totales.
Fuente: elaboración propia

PREGUNTA	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	All
RESPUESTA_ID																	
1	35	24	25	30	32	36	36	39	33	35	47	48	44	27	45	38	574
2	28	29	18	40	7	42	39	41	25	27	37	63	35	24	27	47	529
3	63	69	50	107	39	99	115	140	73	51	107	179	68	37	83	143	1423
4	253	214	203	282	159	282	264	263	234	200	254	270	221	129	167	368	3763
5	808	851	891	728	950	728	733	704	822	874	742	627	819	970	865	591	12703
All	1187	1187	1187	1187	1187	1187	1187	1187	1187	1187	1187	1187	1187	1187	1187	1187	18992

Tabla 10 - Tabla de frecuencia con totales.
Fuente: elaboración propia

PREGUNTA	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	All
RESPUESTA_ID																	
1	0.18	0.13	0.13	0.16	0.17	0.19	0.19	0.21	0.17	0.18	0.25	0.25	0.23	0.14	0.24	0.2	3.02
2	0.15	0.15	0.09	0.21	0.04	0.22	0.21	0.22	0.13	0.14	0.19	0.33	0.18	0.13	0.14	0.25	2.79
3	0.33	0.36	0.26	0.56	0.21	0.52	0.61	0.74	0.38	0.27	0.56	0.94	0.36	0.19	0.44	0.75	7.49
4	1.33	1.13	1.07	1.48	0.84	1.48	1.39	1.38	1.23	1.05	1.34	1.42	1.16	0.68	0.88	1.94	19.81
5	4.25	4.48	4.69	3.83	5	3.83	3.86	3.71	4.33	4.6	3.91	3.3	4.31	5.11	4.55	3.11	66.89
All	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	100

Observar esta información en forma de tablas es útil, sin embargo, se optó por la opción gráfica de los diagramas de dispersión para continuar analizando la relación entre preguntas y respuestas.

4.2.2. Histogramas

La Figura 1 muestra el histograma con las respuestas a las preguntas del constructo 1 “Satisfacción con las actividades del curso”. Se puede observar lo siguiente:

- Hay una inclinación hacia la derecha. La respuesta con más frecuencia fue la 5
- Por otro lado, la respuesta 2 fue la que tuvo menos frecuencia

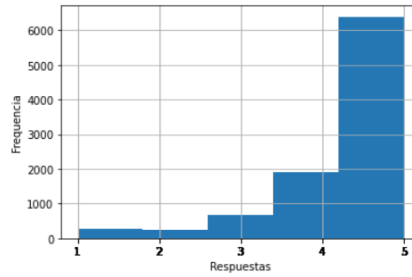


Figura 1 - Histograma Constructo 1.
Fuente: elaboración propia

La Figura 2 muestra el histograma con las respuestas a las preguntas del constructo 2 “Satisfacción con el profesor”. Se puede observar lo siguiente:

- Se tienen resultados muy similares al constructo 1 “Satisfacción con el curso”
 - La respuesta 5 tuvo la mayor frecuencia
 - La respuesta 2 tuvo la menor frecuencia

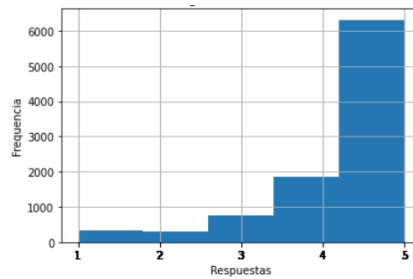


Figura 2 - Histograma Constructo 2.
Fuente: elaboración propia

La Figura 3 muestra el histograma con las respuestas a las preguntas del constructo 3 “Satisfacción en general”. Se puede observar lo siguiente:

- El constructo 3 “Satisfacción en general” es la combinación de las preguntas del constructo 1 “Satisfacción con el curso” y el constructo 2 “Satisfacción con el profesor”. Por lo tanto, se obtuvo que la respuesta con más frecuencia fue la 5 y la menor fue la 2.

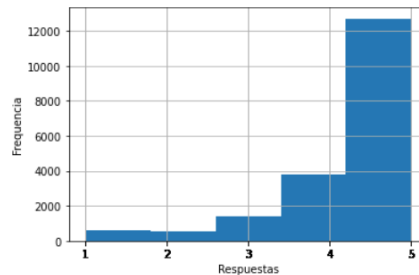


Figura 3 - Histograma Constructo 3.
Fuente: elaboración propia

4.2.3. Estudio explicativo del comportamiento de los constructos en el tiempo

En esta sección se realizará un estudio explicativo con base a series de tiempo y datos por periodo (obtenidos de la fuente de datos previamente analizada), se tienen datos de primaveras y otoños en los años 2016 al 2019 (este último solo hay información de primavera). El objetivo es identificar patrones a través del tiempo, para ello se construyeron diferentes gráficas.

4.2.3.1. Construcción de las series de tiempo

Las gráficas serán construidas con base a los datos obtenidos en la sección 4 “Exploración de los datos”. Además, un conjunto de datos estadísticos fue calculado para la creación de series de tiempo, en la Tabla 11 se puede observar una muestra de los datos obtenidos

Tabla 11 - Series de tiempo con datos estadísticos.
Fuente: elaboración propia

period	construct	mean	std	min	25%	50%	75%	max
P2016	209	4.57	0.80	1.0	4.0	5.0	5.0	5.0
P2016	1017	4.50	0.88	1.0	4.0	5.0	5.0	5.0
P2016	217	4.53	0.84	1.0	4.0	5.0	5.0	5.0
O2016	209	4.45	0.90	1.0	4.0	5.0	5.0	5.0
O2016	1017	4.38	0.97	1.0	4.0	5.0	5.0	5.0

4.2.3.2. Series de tiempo con valores de medias

A continuación, se muestran las figuras obtenidas para las series de tiempo para cada constructo. La Figura 4 muestra la serie de tiempo para las medias del constructo 1 “Satisfacción con el curso”. Por otro lado, la Figura 5 muestra los resultados obtenidos de tendencia, estacionalidad y ruido para estos mismos datos y la Figura 6 muestra la gráfica de medias móviles. Se tiene el mismo conjunto de figuras para la desviación estándar y el valor mínimo.

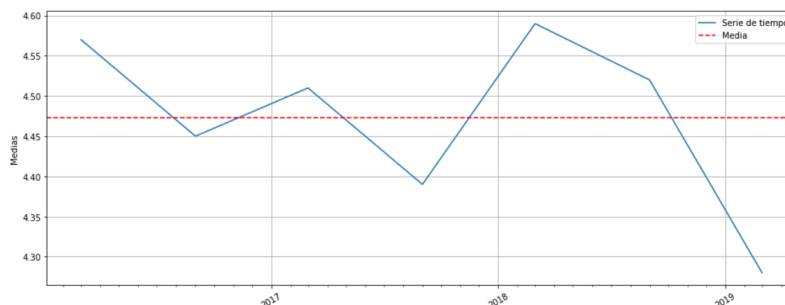


Figura 4 - Serie de tiempo con valores de medias para el Constructo 1.
Fuente: elaboración propia

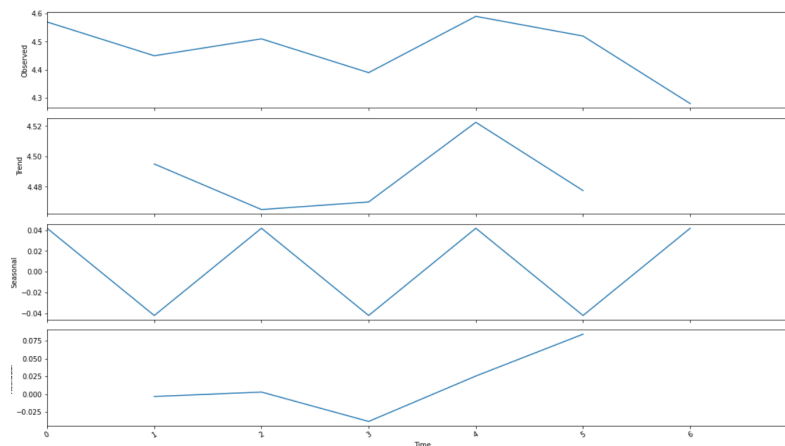


Figura 5 - Tendencia, estacionalidad y ruido, con valores de medias para el Constructo 1.
Fuente: elaboración propia



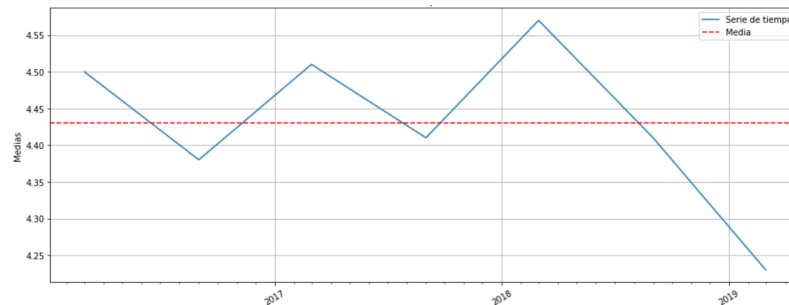
Figura 6 – Medias móviles para el Constructo 1..
Fuente: elaboración propia

Observaciones:

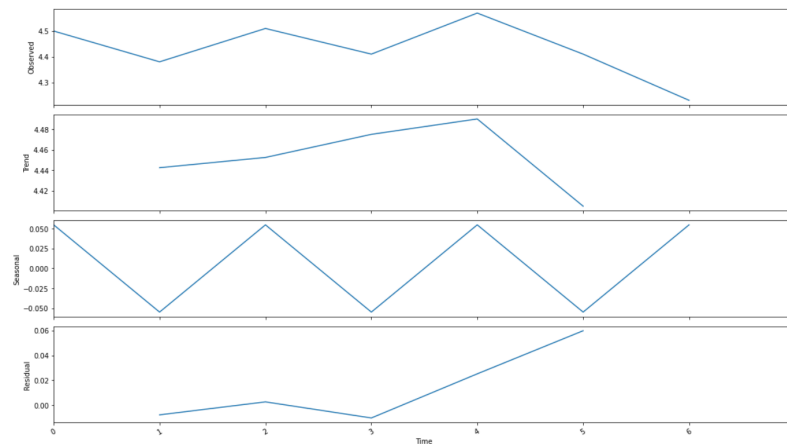
- Los valores obtenidos en las medias de primavera son mayores a las obtenidas en los otoños. Por ejemplo, en primavera 2016 tenemos una media mayor a 4.5 y otoño 2016 tiene un valor menor.
 - Aquí se puede observar un patrón de estacionalidad excepto para primavera 2019
- Mejor y menor valor obtenido en primavera:
 - 2018 tiene el mayor valor en medias con 4.59
 - 2019 tiene el menor valor en medias con 4.28
- Mejor y menor valor obtenido en otoño:
 - 2018 con el mayor valor en medias: 4.39
 - 2017 con el menor valor en medias: 4.52
- Es interesante que las primaveras de 2016, 2017 y 2018 tienen valores mayores a 4.5 pero en el caso de 2019, su valor no alcanzó el 4.3. Se recomienda a la coordinación docente de esta asignatura buscar si las causas estuvieron asociadas al desempeño de algún profesor en específico o a cambios en el programa de estudio o las metodologías. Dado que este estudio no puede hacer un diagnóstico del origen de estas diferencias.

- Entre el 2016 y 2017, se observa que los valores de las medias aumentan y disminuyen, pero a partir del 2018 solo disminuyen
 - Este componente tiene que ver con la tendencia de la serie de tiempo

La Figura 7, Figura 8 y Figura 9 muestran la serie de tiempo con las medias del constructo 2 “Satisfacción con el profesor”, los valores de tendencia, estacionalidad y ruido y, por último, las medias móviles, respectivamente.



Fuente: elaboración propia



Fuente: elaboración propia



Fuente: elaboración propia

Observaciones:

- Los valores obtenidos en las medias de primavera son mayores a las obtenidas en los otoños. Por ejemplo, en primavera 2016 tenemos una media de 4.5 y otoño 2016 tiene un valor menor.
 - Aquí se puede observar un patrón de estacionalidad excepto para primavera 2019
- Mejor y menor valor obtenido en primavera:
 - 2018 tiene el mayor valor en medias con 4.57
 - 2019 tiene el menor valor en medias con 4.23
- Mejor y menor valor obtenido en otoño:
 - 2017 y 2018 tuvieron el mismo valor en medias: 4.41
 - 2016 con el menor valor en medias: 4.38
- Es interesante que las primaveras de 2016, 2017 y 2018 tienen valores igual o mayor a 4.5 pero en el caso de 2019, su valor no alcanzó el 4.3
 - Se obtuvo un resultado similar en el constructo 1 “Satisfacción con el curso”

La Figura 10, Figura 11 y Figura 12 muestran la serie de tiempo con las medias del constructo 3 “Satisfacción en general”, los valores de tendencia, estacionalidad y ruido y, por último, la gráfica de medias móviles, respectivamente.

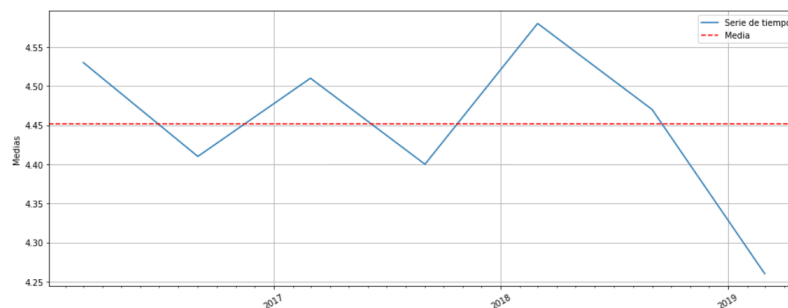


Figura 10 - Serie de tiempo con valores de medias para el Constructo 3.
Fuente: elaboración propia

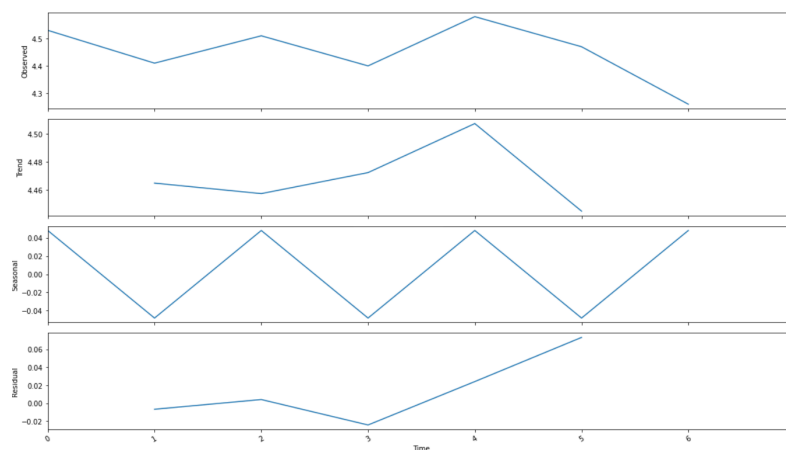


Figura 11 - Tendencia, estacionalidad y ruido con valores de medias para el Constructo 3.
Fuente: elaboración propia

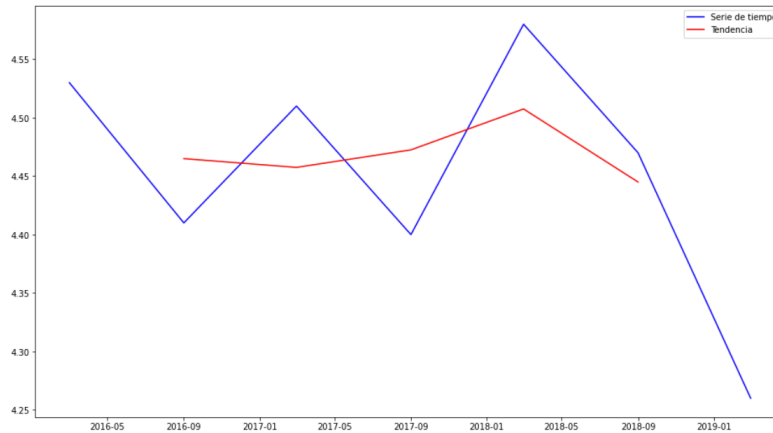


Figura 12 – Medias móviles para el Constructo 3.
Fuente: elaboración propia

Observaciones:

- Los valores obtenidos en las medias de primavera son mayores a las obtenidas en los otoños. Por ejemplo, en primavera 2016 tenemos una media mayor a 4.5 y otoño 2016 tiene un valor menor.
 - Aquí se puede observar un patrón de estacionalidad excepto para primavera 2019
- Mejor y menor valor obtenido en primavera:
 - 2018 tiene el mayor valor en medias con 4.58
 - 2019 tiene el menor valor en medias con 4.26
 - Mismos años que constructos previos
- Mejor y menor valor obtenido en otoño:
 - 2018 con el mayor valor en medias: 4.47
 - 2016 con el menor valor en medias: 4.40
 - Mismos años que el constructo 1
- Es interesante que las primaveras de 2016, 2017 y 2018 tienen valores igual o mayor a 4.5 pero en el caso de 2019, su valor no alcanzó el 4.3
 - Se obtuvo un resultado similar con el constructo 1 “Satisfacción con el curso” y el constructo 2 “Satisfacción con el profesor”
- Entre el 2016 y 2017, se observa que los valores de las medias aumentan y disminuyen, pero a partir del 2018 solo disminuyen
 - Se obtuvo un resultado similar con el constructo 1 “Satisfacción con el curso” y el constructo 2 “Satisfacción con el profesor”

En general, se encontró más similitud entre los resultados obtenidos para los constructos 1 “Satisfacción con el curso” y 3 “Satisfacción en general”. Con base a estos resultados, gran parte de la gente que tomó “Algoritmos y Programación” y contestó la encuesta del IAE entre los periodos 2016 y 2019, estuvieron satisfechos con el curso.

4.2.3.3. Gráficas de cajas y bigotes (Boxplot)

Para las gráficas de cajas y bigotes, se utilizaron los datos por periodo. Cada una de las gráficas despliegan el valor de la media con un punto blanco; además, el cuartil 2 o mediana es representado

con una línea roja. La Figura 13 muestra la gráfica obtenida para el constructo 1 (Satisfacción con las actividades del curso) en los diferentes periodos. Además, la Tabla 12 muestra los datos obtenidos en las series de tiempo que corresponden con la información desplegada en la gráfica.

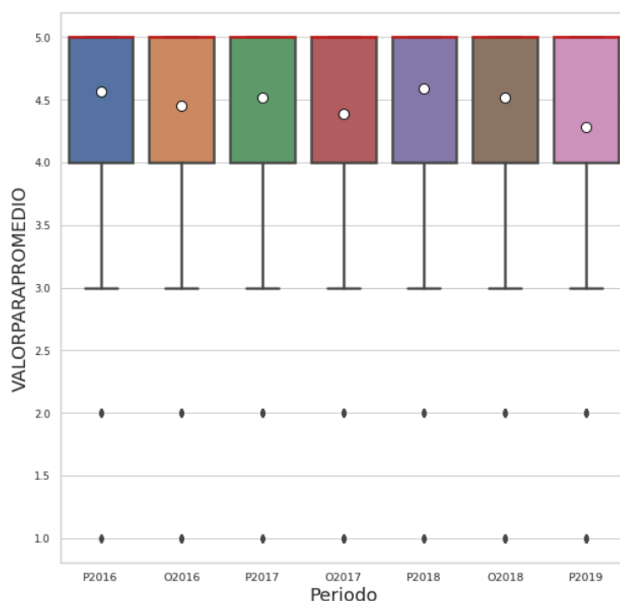


Figura 13 - Boxplot del Constructo 1.
Fuente: elaboración propia

Tabla 12 - Serie del Constructo 1.
Fuente: elaboración propia

	period	mean	std	min	25%	50%	75%	max
0	P2016	4.57	0.80	1.0	4.0	5.0	5.0	5.0
3	O2016	4.45	0.90	1.0	4.0	5.0	5.0	5.0
6	P2017	4.51	0.79	1.0	4.0	5.0	5.0	5.0
9	O2017	4.39	1.10	1.0	4.0	5.0	5.0	5.0
12	P2018	4.59	0.81	1.0	4.0	5.0	5.0	5.0
15	O2018	4.52	0.89	1.0	4.0	5.0	5.0	5.0
18	P2019	4.28	1.09	1.0	4.0	5.0	5.0	5.0

Nótese que en la columna “period” se incluyen los años con un prefijo “P” u “O”, los cuales corresponden a “Primavera” y “Otoño”, respectivamente. Se realizaron los siguientes hallazgos para el constructo 1 “Satisfacción con el curso”:

- Los cuartiles fueron los mismos para todos los periodos
- Las columnas con los cuartiles 2 (50%), 3 (75%) y “max”, tienen el mismo valor: 5; lo cual significa que la mitad de la población calificó con la máxima puntuación las preguntas de este constructo
- Los periodos tienen outliers con los valores 1 y 2, lo cual significa que éstas son respuestas atípicas

Con base en estos resultados, el 75% de los alumnos que respondieron la encuesta de la materia “Algoritmos y Programación” entre los años del 2016 y 2019, estuvieron parcial o totalmente satisfechos con las actividades del curso. Por otro lado, la Figura 14 muestra el boxplot obtenido para el constructo 2 “Satisfacción con el desempeño del profesor” y la Tabla 13 permite observar los datos desde otra perspectiva.

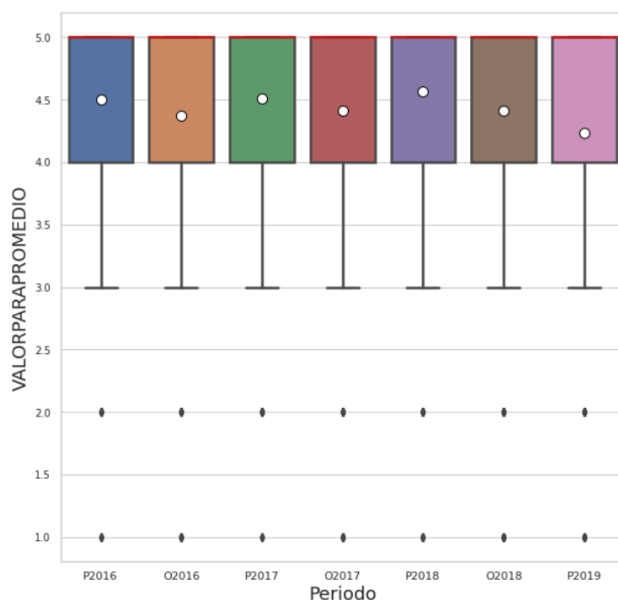


Figura 14 - Boxplot del Constructo 2.
Fuente: elaboración propia

Tabla 13 - Serie del Constructo 2.
Fuente: elaboración propia

	period	mean	std	min	25%	50%	75%	max
1	P2016	4.50	0.88	1.0	4.0	5.0	5.0	5.0
4	O2016	4.38	0.97	1.0	4.0	5.0	5.0	5.0
7	P2017	4.51	0.82	1.0	4.0	5.0	5.0	5.0
10	O2017	4.41	1.10	1.0	4.0	5.0	5.0	5.0
13	P2018	4.57	0.83	1.0	4.0	5.0	5.0	5.0
16	O2018	4.41	0.99	1.0	4.0	5.0	5.0	5.0
19	P2019	4.23	1.19	1.0	4.0	5.0	5.0	5.0

Se puede observar lo siguiente para el Constructo 2 “Satisfacción con el profesor”:

- Los cuartiles, 1, 2, 3 fueron los mismos para todas las temporadas y periodos
- Las columnas con los cuartiles 2 (50%), 3 (75%) y “max” tienen el mismo valor: 5; lo cual significa que gran parte de la gente que respondió, seleccionó ese ítem
- Todas las temporadas tienen *outliers* con los valores 1 y 2, lo cual significa que gran parte de la gente que respondió casi no seleccionó esas opciones como respuesta.

Basado en esos resultados, gran parte de la gente que tomó “Algoritmos y Programación” entre los periodos del 2016 y 2019, estuvieron satisfechos con el desempeño del profesor. Por otro lado, la Figura 15 muestra el *boxplot* para el constructo 3 “Satisfacción general con el curso”; además, la Tabla 14 muestra los valores implementados para construir dicho gráfico.

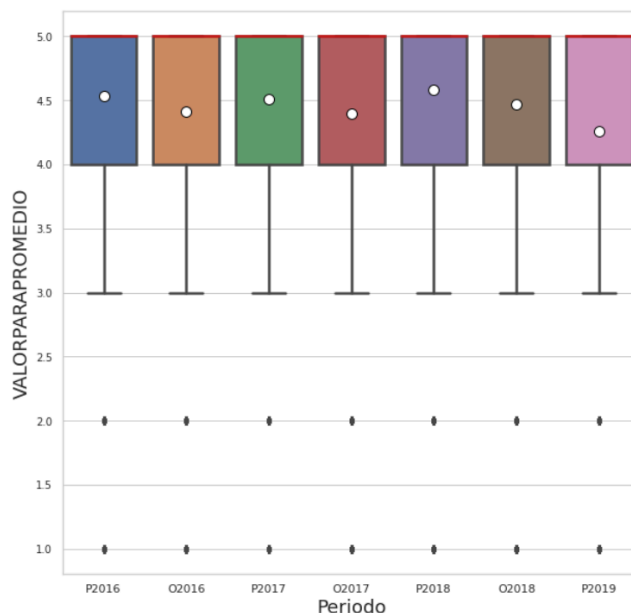


Figura 15 - Boxplot del Constructo 3
Fuente: elaboración propia

Tabla 14 - Series del Constructo 3.
Fuente: elaboración propia

	period	mean	std	min	25%	50%	75%	max
2	P2016	4.53	0.84	1.0	4.0	5.0	5.0	5.0
5	O2016	4.41	0.94	1.0	4.0	5.0	5.0	5.0
8	P2017	4.51	0.81	1.0	4.0	5.0	5.0	5.0
11	O2017	4.40	1.10	1.0	4.0	5.0	5.0	5.0
14	P2018	4.58	0.82	1.0	4.0	5.0	5.0	5.0
17	O2018	4.47	0.94	1.0	4.0	5.0	5.0	5.0
20	P2019	4.26	1.14	1.0	4.0	5.0	5.0	5.0

Se puede observar lo siguiente para el constructo 3 “Satisfacción en general”:

- Los cuartiles, 1, 2, 3 fueron los mismos para todas las temporadas y periodos
 - Resultado similar a los constructos previos
- El cuartil 2 y 3 tienen el mismo valor: 5; lo cual significa que gran parte de la gente que respondió, seleccionó ese ítem
 - Resultado similar a los constructos previos
- Todas las temporadas tienen *outliers* con los valores 1 y 2, lo cual significa que gran parte de la gente que respondió casi no seleccionó esas opciones como respuesta.

- Resultado similar a los constructos previos

4.3. Resultados de la fase de preparación de los datos

Se calculó el coeficiente de correlación Spearman para obtener la relación lineal entre las preguntas de los tres constructos “Satisfacción con el curso”, “Satisfacción con el profesor” y “Satisfacción en general”. El coeficiente de correlación de Spearman se puede usar cuando se tienen variables ordinales [8], por lo tanto, se puede utilizar para esta investigación ya que se tienen ítems tipo Likert cuyas escalas son ordinales.

La Figura 16, Figura 17 y Figura 18 muestran los resultados obtenidos al calcular el coeficiente de correlación Spearman para los tres constructos, se diseñaron mapas de calor para facilitar el análisis de cada uno. Se buscó el conjunto de preguntas con los valores más altos de correlación, se consideró a partir de 0.59 como una correlación importante, dado que el rango de (0–0.5) es considerado como una correlación positiva débil, 0.5 es una correlación positiva moderada y (0.5–1) es una correlación positiva fuerte [38].

En el mapa de calor correspondiente al constructo 1 - “satisfacción con las actividades del curso”, se encontraron correlaciones importantes entre diferentes preguntas. La Tabla 15, Tabla 16 y Tabla 17 incluyen los resultados obtenidos para cada constructo.

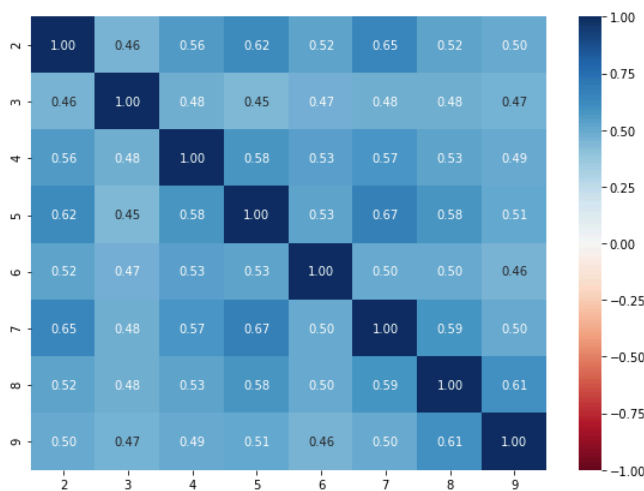


Figura 16 - Mapa de calor del Constructo 1.
Fuente: elaboración propia

Tabla 15 - Correlaciones importantes del Constructo 1 “Satisfacción con el curso”.
Fuente: elaboración propia

Preguntas	Valor obtenido	Observaciones
-----------	----------------	---------------

Pregunta 2: Me han ayudado a identificar los posibles usos o aplicaciones que puede tener lo que aprendo Pregunta 5: Me han ayudado a reflexionar acerca de los aprendizajes que he ido logrando	0.62	Preguntas que piden al estudiante puntuar el conocimiento adquirido en el curso
Pregunta 2: Me han ayudado a identificar los posibles usos o aplicaciones que puede tener lo que aprendo Pregunta 7: Me han ayudado a relacionar con el contexto actual lo que aprendo en clase	0.65	Preguntas que piden al estudiante, puntuar el conocimiento adquirido en el curso
Pregunta 5: Me han ayudado a reflexionar acerca de los aprendizajes que he ido logrando Pregunta 7: Me han ayudado a relacionar con el contexto actual lo que aprendo en clase	0.67	Preguntas que piden al estudiante, puntuar el conocimiento adquirido en el curso Confirma la relación entre las preguntas 2, 5, 7
Pregunta 7: Me han ayudado a relacionar con el contexto actual lo que aprendo en clase Pregunta 8: Han requerido mi participación en situaciones que me permiten aprender a respetar distintas formas de ser y de pensar	0.59	No encuentro mucha relación entre estas preguntas. Se obtuvo justo en el límite de aceptación
Pregunta 8: Han requerido mi participación en situaciones que me permiten aprender a respetar distintas formas de ser y de pensar Pregunta 9: Han requerido que consulte información en diversas fuentes	0.61	Preguntas que piden al estudiante cambiar su forma de realizar actividades

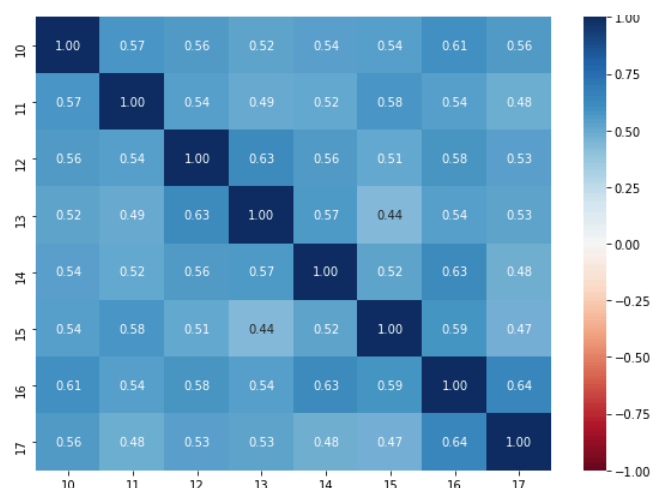


Figura 17 - Mapa de calor del Construto 2.
Fuente: elaboración propia

Tabla 16 - Correlaciones importantes del Construto 2 “Satisfacción con el desempeño del profesor”.
Fuente: elaboración propia

Pregunta	Valor Obtenido	Observaciones
Pregunta 10: Me ha ayudado a tener un panorama claro acerca de lo que aprendería	0.61	Para algunas personas la pregunta 10 puede afectar la decisión de la pregunta 16
Pregunta 16: Recomendando plenamente a este profesor		
Pregunta 12: Me ha ayudado a saber cuáles son los aspectos que debo mejorar	0.63	Ambas preguntas apuntan a que el profesor logró hacer reflexionar y mejorar al estudiante
Pregunta 13: Me ha orientado para mejorar mi capacidad de expresar mis ideas de manera oral y/o escrita		
Pregunta 14: Ha tomado en cuenta la opinión del grupo para mejorar sus clases	0.63	La pregunta 14 puede afectar la decisión de la pregunta 16.
Pregunta 16: Recomendando plenamente a este profesor		
Pregunta 16: Recomendando plenamente a este profesor	0.64	Si se recomienda al profesor puede estar relacionado al nivel de satisfacción con los aprendizajes
Pregunta 17:Cuál es tu nivel de satisfacción con los aprendizajes que has obtenido hasta el momento en esta asignatura		

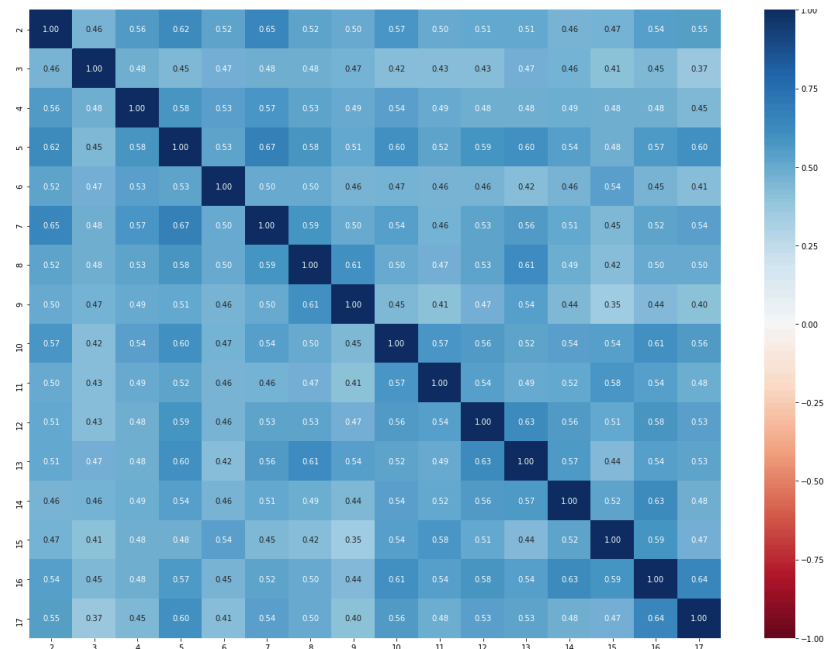


Figura 18 - Mapa de calor del Constructo 3.
Fuente: elaboración propia

Tabla 17 - Correlaciones importantes del Constructo 3 “Satisfacción general con el curso”.
Fuente: elaboración propia

Constructo 1	Constructo 2	Valor obtenido	Observaciones
Pregunta 5: Me han ayudado a reflexionar acerca de los aprendizajes que he ido logrando	Pregunta 10: Me ha ayudado a tener un panorama claro acerca de lo que aprendería	0.60	La pregunta 5 pide reflexión sobre aprendizajes obtenidos, mientras que las preguntas 5, 12, 13 y 17 piden identificar aprendizaje y áreas de oportunidad
	Pregunta 12: Me ha ayudado a saber cuáles son los aspectos que debo mejorar	0.59	
	Pregunta 13: Me ha orientado para mejorar mi capacidad de expresar mis ideas de manera oral y/o escrita	0.60	
	Pregunta 17:Cuál es tu nivel de satisfacción con los aprendizajes que has obtenido hasta el momento en esta asignatura	0.60	

Pregunta 8: Han requerido mi participación en situaciones que me permiten aprender a respetar distintas formas de ser y de pensar	Pregunta 13: Me ha orientado para mejorar mi capacidad de expresar mis ideas de manera oral y/o escrita	0.61	No veo mucha relación entre estas preguntas
---	---	------	---

Con base a estos resultados se puede concluir lo siguiente:

- Todos los resultados fueron positivos
- No todas las preguntas tienen fuertes correlaciones con otras
- La correlación más alta fue entre las siguientes preguntas con un valor de 0.67:
 - Pregunta 5: Me han ayudado a reflexionar acerca de los aprendizajes que he ido logrando
 - Pregunta 7: Me han ayudado a relacionar con el contexto actual lo que aprendo en clase
- La pregunta 5 es la que presenta más correlaciones con otras preguntas de diferentes bloques
- La correlación más débil fue entre las preguntas 3 y 17 con un valor de 0.37
 - Pregunta 3: Han requerido que colabore con otros compañeros
 - Pregunta 7: Me han ayudado a relacionar con el contexto actual lo que aprendo en clase
- Es interesante que exista una correlación entre las preguntas 16 y 17 debido a que relacionan la recomendación del profesor con el nivel de satisfacción de la asignatura
 - Pregunta 16: Recomendando plenamente a este profesor
 - Pregunta 17: Cuál es tu nivel de satisfacción con los aprendizajes que has obtenido hasta el momento en esta asignatura

4.3.1. Mapa de clúster o dendograma

El análisis de clúster se puede utilizar para agrupar ítems y se encuentra estrechamente relacionado a los mapas de calor creados. El mapa de clúster se genera con base la información incluida en la matriz de correlación, permite que las correlaciones sean informativas en cuanto al análisis de las respuestas de las preguntas, debido a que implementan una agrupación jerárquica para agrupar las preguntas en función de su relación [39]. Por lo tanto, el mapa de clúster permitirá visualizar la información segmentada en grupos, lo que facilitará el análisis.

El eje horizontal del dendograma representa la distancia o la diferencia entre los grupos, mientras que el eje vertical representa los objetos y grupos. Se obtuvieron tres mapas de clúster. La Figura

19, Figura 20 y Figura 21 muestran los resultados obtenidos para los tres constructos. Para el constructo 1 “Satisfacción con las actividades del curso”, se formaron los siguientes clústeres:

- Clúster 1:
 - Pregunta 3: Han requerido que colabore con otros compañeros
 - Pregunta 6: Han requerido que aprenda mediante el uso de tecnologías de información y comunicación
- Clúster 2:
 - Pregunta 2: Me han ayudado a identificar los posibles usos o aplicaciones que puede tener lo que aprendo
 - Pregunta 4: Han promovido que aprenda a partir de mis propias acciones
 - Pregunta 5: Me han ayudado a reflexionar acerca de los aprendizajes que he ido logrando
 - Pregunta 7: Me han ayudado a relacionar con el contexto actual lo que aprendo en clase
 - Pregunta 8: Han requerido mi participación en situaciones que me permiten aprender a respetar distintas formas de ser y de pensar
 - Pregunta 9: Han requerido que consulte información en diversas fuentes

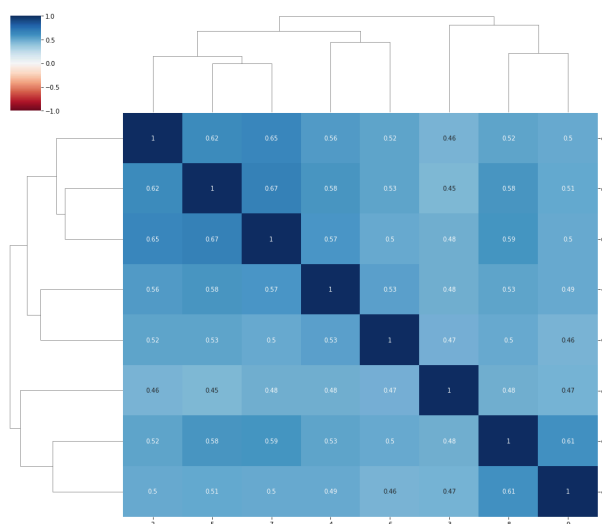


Figura 19 - Clústeres del Constructo 1.
Fuente: elaboración propia

Nótese que el clúster 2 incluye todas las preguntas con correlaciones fuertes encontradas en el constructo 1. Es interesante que además incluye a la pregunta 4 porque no hubo correlaciones fuertes con ésta. Por lo tanto, el clúster 1 contiene preguntas con correlaciones débiles. Por otro lado, para el constructo 2 “Satisfacción con el desempeño del profesor”, se crearon los siguientes clústeres:

- Clúster 1:
 - Pregunta 11: Me ha ayudado a entender con claridad los criterios con los que sería evaluada(o)
 - Pregunta 15: Ha mostrado dominio de los conocimientos y habilidades propios de esta asignatura

- Clúster 2:
 - Pregunta 10: Me ha ayudado a tener un panorama claro acerca de lo que aprendería
 - Pregunta 12: Me ha ayudado a saber cuáles son los aspectos que debo mejorar
 - Pregunta 13: Me ha orientado para mejorar mi capacidad de expresar mis ideas de manera oral y/o escrita
 - Pregunta 14: Ha tomado en cuenta la opinión del grupo para mejorar sus clases
 - Pregunta 16: Recomendando plenamente a este profesor
 - Pregunta 17:Cuál es tu nivel de satisfacción con los aprendizajes que has obtenido hasta el momento en esta asignatura

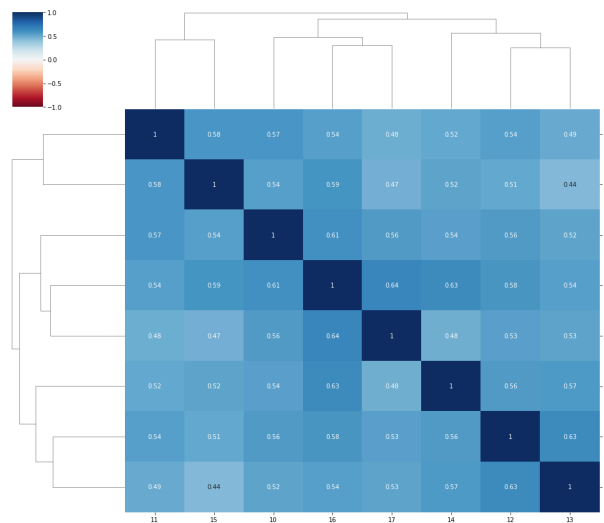


Figura 20 - Clústeres del Constructo 2.
Fuente: elaboración propia

Nótese que el clúster 1 incluye todas las preguntas con correlaciones débiles mientras que el 2 incluye a todas las que tuvieron correlaciones fuertes. Esto es diferente a los resultados obtenidos con el constructo 1, en donde el clúster 2 incluyó una pregunta que no tuvo correlaciones fuertes con las demás. Finalmente, para el constructo 3 “Satisfacción general con el curso” se obtuvieron los siguientes clústeres

- Clúster con preguntas del Constructo 1
 - Preguntas 2, 3, 4, 5, 7, 7, 8, 9
- Clúster con preguntas del Constructo 2
 - Preguntas 10, 11, 12, 13, 14, 15, 16, 17

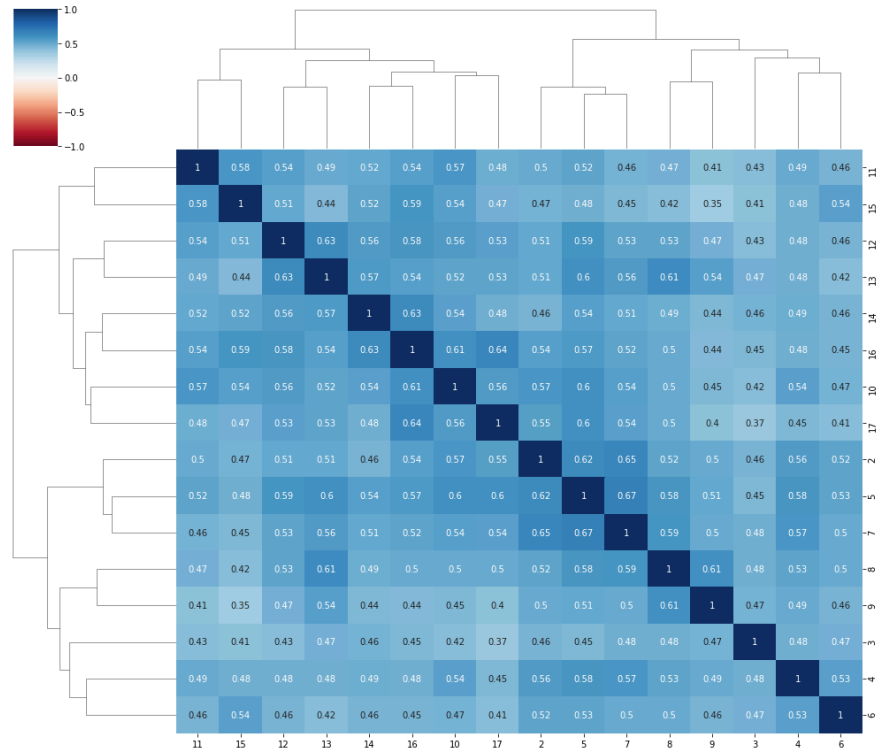


Figura 21 - Clústeres del Constructo 3.
Fuente: elaboración propia

Nótese que los dos clústeres obtenidos corresponden a las preguntas de los constructos 1 y 2. El primer clúster con preguntas de la 2 – 17, el 2do con las preguntas 10 – 17. Es interesante que las preguntas de los constructos terminaran en clústeres diferentes.

4.4. Resultados de la fase de modelado

Se calculó la prueba de Kaiser-Meyer-Olkin (KMO) para evaluar si el conjunto de datos es adecuado para realizar el análisis factorial, el análisis se realizó sobre el constructo 3 “Satisfacción en general” con las preguntas 2-17. La prueba KMO se calculó usando la librería `factor_analyzer`, obteniendo como resultado 0.9, el cual supera el valor mínimo apropiado para el análisis factorial de 0.5 [11].

Se realizó también la prueba de Bartlett para confirmar el resultado obtenido con KMO y determinar si el conjunto de datos es adecuado para realizar el análisis factorial. La prueba se realizó sobre el constructo 3 “Satisfacción en general” con las preguntas 2-17, utilizando la librería `factor_analyzer` se obtuvo un valor p de 0.0, este resultado es significativo e indica que la matriz de correlación no es una matriz de identidad, por lo tanto, los valores son adecuados para el análisis factorial [12].

4.4.1. Análisis factorial exploratorio

Se realizó el EFA con diferentes rotaciones ortogonales y oblicuas para comparar los resultados. El método de ajuste utilizado fue MINRES, el cual proporciona buenos resultados con la matriz de correlación. La Tabla 18 muestra los valores obtenidos. Se consideró 0.6 como el límite inferior aceptable para considerar las cargas obtenidas durante el cálculo de EFA, se seleccionó este valor más conservador que 0.4 ya que es demasiado liberal para esta investigación.

Tabla 18 - Análisis factorial con diferentes rotaciones.
Fuente: elaboración propia

Rotación	Tipo de Rotación	Límite	Factor 1	Factor 2	Preguntas Faltantes
varimax	ortogonal	0.6	2, 4, 5, 7, 8, 9, 13	10, 11, 12, 14, 15, 16, 17	3, 6
promax	oblicua	0.6	5, 7, 8, 9, 13	10, 11, 14, 15, 16, 17	2, 3, 4, 6, 12
oblimin	oblicua	0.6	2-17		
quartimax	ortogonal	0.6	2-17		
equamax	ortogonal	0.6	2-17		

Con la rotación varimax se obtuvieron dos factores, con el primer factor se puede crear un nuevo constructo, todas las preguntas que contiene pertenecen al constructo 1 “Satisfacción con el curso”, solo una pregunta al constructo 2 “Satisfacción con el profesor” y dos preguntas del constructo 1 no fueron incluidas (3, 6). Se puede solicitar a un grupo de expertos asignar el nombre al nuevo constructo con base las siguientes preguntas, las cuales fueron obtenidas con el factor 1 del EFA, la Figura 22 muestra el mapa de clúster del EFA con rotación varimax.

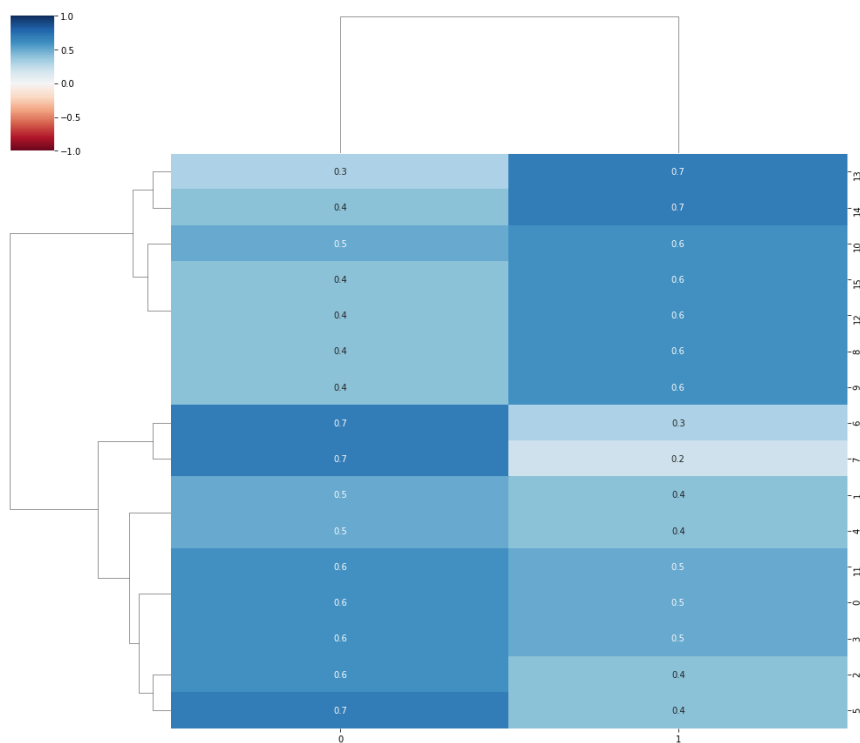


Figura 22 - Mapa de clúster de EFA con rotación varimax.
Fuente: elaboración propia

- Pregunta 2: Me han ayudado a identificar los posibles usos o aplicaciones que puede tener lo que aprendo
- Pregunta 4: Han promovido que aprenda a partir de mis propias acciones
- Pregunta 5: Me han ayudado a reflexionar acerca de los aprendizajes que he ido logrando
- Pregunta 7: Me han ayudado a relacionar con el contexto actual lo que aprendo en clase
- Pregunta 8: Han requerido mi participación en situaciones que me permiten aprender a respetar distintas formas de ser y de pensar
- Pregunta 9: Han requerido que consulte información en diversas fuentes
- Pregunta 13: Me ha orientado para mejorar mi capacidad de expresar mis ideas de manera oral y/o escrita

Con el segundo factor obtenido del EFA y la rotación varimax, se puede crear un nuevo constructo, con preguntas del constructo dos “Satisfacción con el profesor” sin considerar la pregunta 13 “Me ha orientado para mejorar mi capacidad de expresar mis ideas de manera oral y/o escrita”. Con este nuevo listado de preguntas, un grupo de expertos podría asignar un nombre a este nuevo constructo. la Figura 22 muestra el mapa de clúster del EFA con rotación varimax.

- Pregunta 10: Me ha ayudado a tener un panorama claro acerca de lo que aprendería
- Pregunta 11: Me ha ayudado a entender con claridad los criterios con los que sería evaluada(o)
- Pregunta 12: Me ha ayudado a saber cuáles son los aspectos que debo mejorar
- Pregunta 14: Ha tomado en cuenta la opinión del grupo para mejorar sus clases
- Pregunta 15: Ha mostrado dominio de los conocimientos y habilidades propios de esta asignatura
- Pregunta 16: Recomendando plenamente a este profesor
- Pregunta 17: Cuál es tu nivel de satisfacción con los aprendizajes que has obtenido hasta el momento en esta asignatura

Nótese que las siguientes preguntas no fueron consideradas en ninguno de los dos factores, ambas pertenecen al constructo 1 y están relacionadas al trabajo en equipo y uso de tecnologías de la información.

- Pregunta 3: Han requerido que colabore con otros compañeros
- Pregunta 6: Han requerido que aprenda mediante el uso de tecnologías de información y comunicación

Después, se realizó el EFA con la rotación promax obteniendo como resultado dos factores. Con el primer factor se puede crear un nuevo constructo, incluye cuatro preguntas del constructo 1 “Satisfacción con el curso” y una de “Satisfacción con el profesor”. Se puede pedir a un grupo de expertos asignar un nombre al nuevo constructo con base al nuevo listado de preguntas: 5, 7, 8, 9 y 13. El segundo factor obtenido está compuesto por preguntas de “Satisfacción con el profesor”, un grupo de expertos podría asignar un nombre al nuevo constructo con base al nuevo listado de preguntas: 10, 11, 14, 15, 16 y 17. Nótese que las preguntas 2, 3, 4, 6 y 12 no fueron incluidas en estos constructos, es la misma lista obtenida con la rotación varimax, pero con más preguntas del

constructo 1 y una del constructo 2. La Figura 23 muestra el mapa de clúster del EFA con rotación promax.

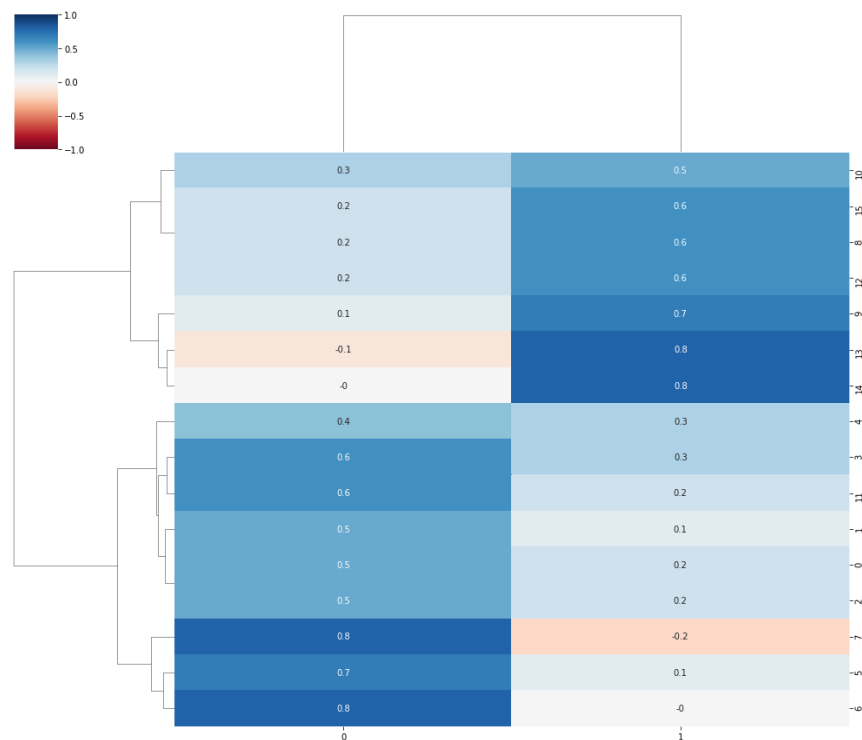


Figura 23 - Mapa de clúster de EFA con rotación promax.
Fuente: elaboración propia

Finalmente, el EFA con las rotaciones oblimin, quartimax y equamax se obtuvo un solo factor con las preguntas 2-17, es decir, las preguntas del constructo 3 “Satisfacción en general”. La Figura 24 muestra el mapa de clúster de EFA sin rotación, los mapas de clúster con rotaciones oblimin, quartimax y equamax proporcionaron el mismo resultado.

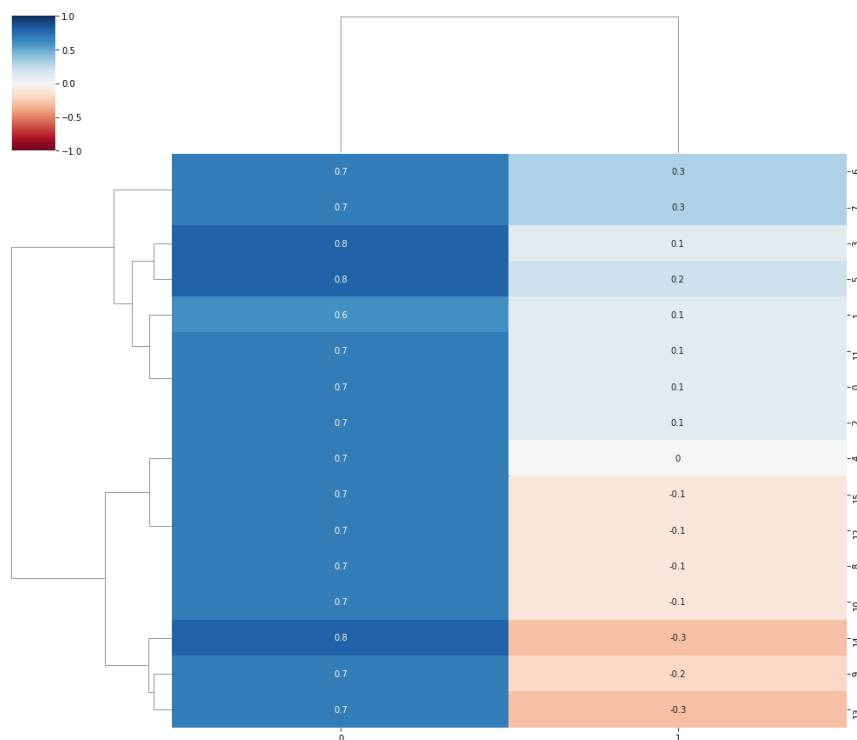


Figura 24 - Mapa de clúster del EFA sin rotación.
Fuente: elaboración propia

4.4.2. Análisis factorial confirmatorio (CFA)

4.4.2.1. Prueba de fiabilidad con el alfa de Cronbach

Para asegurar que se obtuvo una buena escala se requiere la implementación de algunas medidas de calidad, se debe probar su validez y fiabilidad. El coeficiente del alfa de Cronbach es la medida más popular para la confiabilidad interna, introducida por Guttman (1945) [18] y popularizada por Cronbach (1951) [19].

Se obtuvo un alfa de 0.92 con el constructo 1 “Satisfacción con el curso”, un alfa de 0.93 con el constructo 2 “Satisfacción con el profesor” y un alfa de 0.96 “Satisfacción en general”. Para los tres constructos, los valores de alfa indican una alta confiabilidad del instrumento del cuestionario y consistencias internas de las escalas tipo Likert incluidas en el IAE.

5. CONCLUSIONES Y RECOMENDACIONES

A continuación, se presentan las conclusiones obtenidas para diferentes etapas del proceso:

1. Para este trabajo se contó con resultados de 1187 encuestas obtenidas del IAE entre los años 2016 y 2019 sobre la materia de Algoritmos y programación. Para los tres constructos “Satisfacción con el curso”, “Satisfacción con el profesor” y “Satisfacción en general”, se obtuvieron las medias de las primaveras 2016, 2017 y 2018, se encontraron valores mayores a 4.5 pero el año 2019 no alcanzó el 4.3. Se recomienda a la coordinación docente de esta asignatura buscar si las causas estuvieron asociadas al desempeño de algún profesor en específico o cambios en el programa de estudio o las metodologías. Dado que este estudio no puede hacer un diagnóstico del origen de estas diferencias
2. En los tres constructos la respuesta más utilizada fue la opción 5 “Totalmente de acuerdo”, mientras que las menos utilizadas fueron las opciones 1 “Totalmente en desacuerdo” y 2 “Parcialmente en desacuerdo”. Por lo tanto, se puede concluir que más del 50% los estudiantes que tomaron la materia entre 2016 y 2019 estuvieron satisfechos con el curso, el profesor y en general.
3. El EFA con la rotación varimax indicó que se pueden tener dos factores, el primer factor con las preguntas 2, 4, 5, 7, 8, 9 y 13, es decir, las preguntas del constructo 1 “Satisfacción con el curso”, excluyendo las pregunta 3 y 6, las cuales no fueron incluidas en algún constructo, y añadiendo la pregunta 13 del constructo 2 “Satisfacción con el profesor”. El segundo factor compuesto por las preguntas 10, 11, 12, 14, 15, 16 y 17, es decir, las preguntas del constructo 2 excluyendo la pregunta 13 que ya fue incluida en el primer factor. Se puede solicitar a un grupo de expertos asignar el nombre a este nuevo constructo con base a estas preguntas. Por otro lado, el EFA con la rotación promax también indicó que se pueden obtener dos factores, el primer factor con las preguntas 5, 7, 8, 9, 13, es decir, preguntas del constructo 1 excluyendo las preguntas 2, 3, 4 y 6, las cuales no fueron incluidas en algún constructo, y añadiendo la pregunta 13 que pertenece al constructo 2. El segundo factor compuesto por las preguntas 10, 11, 14, 15, 16 y 17, es decir, preguntas del constructo 2 excluyendo 12 y 13, la pregunta 12 no fue incluida en algún constructo y la 13 incluida en el factor 1. Se puede solicitar a un grupo de expertos asignar el nombre a este nuevo constructo con base a estas preguntas. El EFA con las rotaciones oblimin, quartimax y equamax indicaron que se puede tener un único factor con las preguntas del

constructo 1 y 2, es decir, preguntas de la 2 a la 9 que corresponden al constructo 3 “Satisfacción en general”.

4. El CFA con el alfa de Cronbach sobre la hipótesis del constructo 1 “Satisfacción con el curso”, donde se plantea que las preguntas 2 a la 9 permiten medir la satisfacción del estudiante con el curso, se obtuvo 0.92 como resultado. Sobre la hipótesis del constructo 2, en donde se plantea que las preguntas 10 a la 17 permiten medir la satisfacción con el desempeño del profesor, se obtuvo 0.93 como resultado. Finalmente, la hipótesis del constructo 3, donde se plantea que las preguntas 2 a la 17 permiten medir la satisfacción con el curso en general, se obtuvo 0.96 como resultado. Para los tres constructos se obtuvieron resultados excelentes mayores a 0.9, lo que indica alta confiabilidad y consistencia interna para cada constructo.

BIBLIOGRAFÍA

- [1] N. Chrisman, “Beyond Stevens: A revised approach to measurement for geographic information,” *Proceedings of AUTO-CARTO 12*, vol. 12, Jan. 1995.
- [2] L. Cohen, L. Manion, and K. Morrison, “Research Methods in Education,” [http://lst-iiiep.iiep-unesco.org/cgi-bin/wwwi32.exe/\[in=epidoc1.in\]/?t2000=011160/\(100\)](http://lst-iiiep.iiep-unesco.org/cgi-bin/wwwi32.exe/[in=epidoc1.in]/?t2000=011160/(100)), Jan. 2000.
- [3] S. E. Harpe, “How to analyze Likert and other rating scale data,” *Currents in Pharmacy Teaching and Learning*, vol. 7, no. 6, pp. 836–850, Nov. 2015, doi: 10.1016/j.cptl.2015.08.001.
- [4] “Construct | psychology,” *Encyclopedia Britannica*.
<https://www.britannica.com/science/construct> (accessed Sep. 29, 2021).
- [5] S. P. Desselle, “Construction, Implementation, and Analysis of Summated Rating Attitude Scales,” *Am J Pharm Educ*, vol. 69, no. 5, p. 97, Sep. 2005, doi: 10.5688/aj690597.
- [6] “Likert Scales: Dispelling the Confusion.” <http://www.john-uebersax.com/stat/likert.htm> (accessed Nov. 16, 2020).
- [7] L. Restrepo and J. González, “De Pearson a Spearman,” *Revista Colombiana de Ciencias Pecuarías*, vol. 20, pp. 183–192, Jun. 2007.
- [8] P. Morales and L. Rodríguez, “APLICACIÓN DE LOS COEFICIENTES CORRELACIÓN DE KENDALL Y SPEARMAN (Application of the Kendall correlation and Spearman coefficients),” p. 8.
- [9] B. Williams *et al.*, “EDUCATION Exploratory factor analysis: A five-step guide for novices.”
- [10] V. Costa and R. Sarmento, “Confirmatory Factor Analysis,” p. 39.
- [11] K. Chaiwut, W. Rueangsirarak, and R. Chaisricharoen, “Factor analysis on student loan consideration in higher education level,” in *2017 International Conference on Digital Arts, Media and Technology (ICDAMT)*, Mar. 2017, pp. 296–301. doi: 10.1109/ICDAMT.2017.7904980.
- [12] J. Niu *et al.*, “A comparative study on application of data mining technique in human shape clustering: Principal component analysis vs. Factor analysis,” in *2010 5th IEEE Conference on Industrial Electronics and Applications*, Jun. 2010, pp. 2014–2018. doi: 10.1109/ICIEA.2010.5515577.
- [13] K. G. Joreskog, “Factor Analysis by MINRES,” p. 13.
- [14] “Factor rotation > Factor analysis (FA) > Statistical Reference Guide | Analyse-it® 5.65 documentation.” <https://analyse-it.com/docs/user-guide/multivariate/factor-rotation> (accessed Feb. 20, 2022).
- [15] “QUESTIONS AND ANSWERS.”
http://www.harryganzeboom.nl/teaching/factoranalysis_qa.htm (accessed Mar. 22, 2022).
- [16] “Factor Analysis: Varimax Rotation,” in *The SAGE Encyclopedia of Communication Research Methods*, 2455 Teller Road, Thousand Oaks California 91320: SAGE Publications, Inc, 2017. doi: 10.4135/9781483381411.n191.
- [17] H. Abdi, “Factor Rotations in Factor Analyses,” p. 8.
- [18] L. Guttman, “A basis for analyzing test-retest reliability,” *Psychometrika*, vol. 10, no. 4, pp. 255–282, Dec. 1945, doi: 10.1007/BF02288892.
- [19] L. J. Cronbach, “Coefficient alpha and the internal structure of tests,” p. 38.
- [20] F. M. Lord, M. R. Novick, and A. Birnbaum, *Statistical theories of mental test scores*. Oxford, England: Addison-Wesley, 1968.

- [21] S. Contreras Espinoza and F. Novoa-Muñoz, “Ventajas del alfa ordinal respecto al alfa de Cronbach ilustradas con la encuesta AUDIT-OMS,” *Rev Panam Salud Publica*, pp. 1–6, 2018, doi: 10.26633/RPSP.2018.65.
- [22] “CRISP-DM,” *Data Science Process Alliance*. <https://www.datascience-pm.com/crisp-dm-2/> (accessed Apr. 03, 2022).
- [23] S. S. Stevens, “On the Theory of Scales of Measurement,” *Science, New Series*, vol. 103, no. 2684, pp. 677–680, 1946.
- [24] R. P. Chalmers, “On Misconceptions and the Limited Usefulness of Ordinal Alpha,” *Educational and Psychological Measurement*, vol. 78, no. 6, pp. 1056–1071, Dec. 2018, doi: 10.1177/0013164417727036.
- [25] R. Krishnan and K. W. Loon, “The Effects of Job Satisfaction and Work-Life Balance on Employee Task Performance,” *IJARBSS*, vol. 8, no. 3, p. Pages 652-662, Apr. 2018, doi: 10.6007/IJARBSS/v8-i3/3956.
- [26] A. A. Olaniyi, “Application of Likert Scale’s Type and Cronbach’s Alpha Analysis in an Airport Perception Study,” p. 5, 2019.
- [27] “Project Jupyter.” <https://jupyter.org> (accessed Apr. 26, 2022).
- [28] “Google Colaboratory.” https://colab.research.google.com/?utm_source=scs-index (accessed Apr. 26, 2022).
- [29] “sklearn.decomposition.FactorAnalysis — scikit-learn 0.24.1 documentation.” <https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.FactorAnalysis.html> (accessed Mar. 31, 2021).
- [30] “What is NumPy? — NumPy v1.21 Manual.” <https://numpy.org/doc/stable/user/whatisnumpy.html> (accessed Oct. 04, 2021).
- [31] “pandas documentation — pandas 1.3.3 documentation.” <https://pandas.pydata.org/docs/> (accessed Oct. 04, 2021).
- [32] “Getting Started with Plotly.” <https://plotly.com/python/getting-started/> (accessed Oct. 04, 2021).
- [33] R. Vallat, “Pingouin: statistics in Python,” *JOSS*, vol. 3, no. 31, p. 1026, Nov. 2018, doi: 10.21105/joss.01026.
- [34] “scikit-learn: machine learning in Python — scikit-learn 1.0 documentation.” <https://scikit-learn.org/stable/index.html> (accessed Oct. 04, 2021).
- [35] “Matplotlib — Visualization with Python.” <https://matplotlib.org/> (accessed Feb. 13, 2022).
- [36] M. L. Waskom, “seaborn: statistical data visualization,” *Journal of Open Source Software*, vol. 6, no. 60, p. 3021, Apr. 2021, doi: 10.21105/joss.03021.
- [37] “Introduction — statsmodels.” <https://www.statsmodels.org/stable/index.html> (accessed Feb. 13, 2022).
- [38] “¿Qué es una correlación? ... y herramientas de análisis de datos,” *INCIBE-CERT*, Jul. 09, 2015. <https://www.incibe-cert.es/blog/correlacion-herramientas-analisis-datos> (accessed Mar. 31, 2021).
- [39] Y. Holtz, “Dendrogram,” *The Python Graph Gallery*. <https://www.python-graph-gallery.com/dendrogram/> (accessed Mar. 31, 2021).