

**PERFILES DEL EMPLEO DE LOS GRADUADOS DEL 2008 DE LA
UNIVERSIDAD PONTIFICIA BOLIVARIANA DE UN AÑO A DOS AÑOS
DESPUÉS DEL GRADO: UN ANÁLISIS DE CLASIFICACIÓN**

DIANA LICETH GARCÉS PORTILLA

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICAS
ESPECIALIZACIÓN EN ESTADÍSTICA
BUCARAMANGA**

2012

**PERFILES DEL EMPLEO DE LOS GRADUADOS DEL 2008 DE LA
UNIVERSIDAD PONTIFICIA BOLIVARIANA DE UN AÑO A DOS AÑOS
DESPUÉS DEL GRADO: UN ANÁLISIS DE CLASIFICACIÓN**

DIANA LICETH GARCÉS PORTILLA

Trabajo de grado para optar por el título de Especialista en Estadística

Director

GERMÁN MORENO ARENAS

Doctor en Estadística

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE CIENCIAS

ESCUELA DE MATEMÁTICAS

ESPECIALIZACIÓN EN ESTADÍSTICA

BUCARAMANGA

2012

CONTENIDO

	Pág.
1. INTRODUCCIÓN	12
2. MARCO DE REFERENCIA	18
2.1 SUPUESTOS DEL ANÁLISIS DE CLASIFICACIÓN	18
2.1.1 Representatividad de la muestra.	18
2.1.2 Multicolinealidad.	18
2.2 MEDIDA DE SIMILARIDAD	23
2.3 LA FORMACIÓN DE LOS CONGLOMERADOS	25
2.4 NÚMERO DE CONGLOMERADOS	28
2.5 INTERPRETACIÓN DE LOS CONGLOMERADOS	28
2.6 LA VALIDEZ DE LA SOLUCIÓN DE CLASIFICACIÓN	29
3. RESULTADOS	31
3.1 SUPUESTOS DEL ANÁLISIS DE CLASIFICACIÓN	31
3.1.1 Representatividad de la muestra.	31
3.1.2 Multicolinealidad.	31
3.2 FORMACIÓN DE CONGLOMERADOS	37
3.3 VALIDACIÓN DE LA SOLUCIÓN DE TRES CONGLOMERADOS	43
3.4 INTERPRETACIÓN DE LA SOLUCIÓN DE TRES CONGLOMERADOS	45
3.5 TIPOLOGÍA DE LA SOLUCIÓN DE TRES CONGLOMERADOS	50
3.6 OBTENCIÓN DE PERFILES DE LOS TRES CONGLOMERADOS SOBRE VARIABLES ADICIONALES	55
CONCLUSIONES	58
RECOMENDACIONES	60
BILIOGRAFÍA	61
ANEXOS	63

LISTA DE GRÁFICAS

	Pág.
Gráfica 1. Dendograma por el método Ward	40
Gráfica 2. Media de los valores del componente de regresión 1	46
Gráfica 3. Media de los valores del componente de regresión 2	47
Gráfica 4. Media de los valores del componente de regresión 3	48
Gráfica 5. Gráfica de individuos en conglomerados	49

LISTA DE TABLAS

	Pág.
Tabla 1. Variables relacionadas con la situación laboral	15
Tabla 2. Matriz de correlación	32
Tabla 3. KMO y Prueba de Bartlett	33
Tabla 4. Matriz de correlación anti-imagen	33
Tabla 5. Varianza total explicada por los componentes principales	34
Tabla 6. Matriz de componentes principales no rotada	35
Tabla 7. Matriz de componentes principales rotada (Rotación varimax)	35
Tabla 8. Matriz de correlaciones reproducidas	36
Tabla 9. Esquema de aglomeración por el método Ward	37
Tabla 10. Coeficientes de aglomeración	41
Tabla 11. Prueba T para la igualdad de medias componente principal 1 entre dos conglomerados	42
Tabla 12. Prueba T para la igualdad de medias componente principal 2 entre dos conglomerados	42
Tabla 13. Prueba U de Mann-Whitney para la igualdad de medias componente principal 3 entre dos conglomerados	42
Tabla 14. Anova: Solución tres conglomerados	42
Tabla 15. Pertenencia a los conglomerados	44
Tabla 16. Anova Método Umbral Paralelo	45
Tabla 17. Anova Método Ward	45
Tabla 18. Comparaciones múltiples – Diferencia Mínima Significativa (DMS)	46
Tabla 19. Comparaciones múltiples – Diferencia Mínima Significativa (DMS)	47
Tabla 20. Comparaciones múltiples – Diferencia Mínima Significativa (DMS)	49
Tabla 21. Estadísticos descriptivos de los tres conglomerados	50
Tabla 22. Estadísticos descriptivos Salario mensual	52
Tabla 23. Tabla de contingencia variable A2 vs conglomerados	53

Tabla 24. Tabla de contingencia variable A3 vs conglomerados	53
Tabla 25. Tabla de contingencia variable A5 vs conglomerados	53
Tabla 26. Tabla de contingencia variable A6 vs conglomerados	54
Tabla 27. Tabla de contingencia variable A7 vs conglomerados	54
Tabla 28. Tabla de contingencia variable A8 vs conglomerados	54
Tabla 29. Tabla de contingencia variable A8 vs conglomerados	55
Tabla 30. Tabla de contingencia variable género vs conglomerados	55
Tabla 31. Tabla de contingencia variable nivel de responsabilidad en la definición de los objetivos de la empresa vs conglomerados	56
Tabla 32. Tabla de contingencia variable contrato vs conglomerados	56
Tabla 33. Tabla de contingencia variable escuela vs conglomerados	57
Tabla 34. Tabla de contingencia variable facultad vs conglomerados	57
Tabla 35. Tabla de contingencia variable tamaño empresa vs conglomerados	57

LISTA DE ANEXOS

	Pág.
ANEXO A. Análisis descriptivo univariado	63
ANEXO B. Directrices para la identificación de las cargas factoriales significativas basadas en el tamaño muestral, con nivel de significación del 0,05 y un nivel de potencia del 80%	70
ANEXO C. Análisis descriptivo bivariado	71

RESUMEN

Título: Perfiles del empleo de los graduados del 2008 de la Universidad Pontificia Bolivariana de un año a dos años después del grado: Un análisis de clasificación.

Autor: Diana Liceth Garcés Portilla**

Palabras Claves: Graduados, características del trabajo, análisis factorial y de clasificación.

Descripción:

El presente estudio busca agrupar a los graduados del 2008 de la Universidad Pontificia Bolivariana de un año a dos años después del grado en grupos con características del empleo similares y a la vez que sean lo más disímiles entre sí.

Esta reducción de los graduados en grupos facilita la interpretación de los datos, dado que se establecen perfiles del empleo. Para llegar a la definición de los perfiles, se cumplieron varias etapas, entre ellas, la reducción de variables correlacionadas por medio del análisis factorial para ingresarlas en el análisis de clasificación. Luego se determinó la medida de similitud, la cual guarda relación con el tipo de variables de estudio. En este caso se hace uso de la distancia euclidiana al cuadrado, ya que unas variables son cuantitativas y otras por el escalamiento realizado, en forma de escala de Likert, permite ser tratadas como medidas métricas.

Determinada la medida de similitud de distancia, se aplicó el método de formación de conglomerados Ward, ya que es un procedimiento que se basa en la medida de distancia euclidiana al cuadrado. Establecida la estructura de los conglomerados, para definir cuantos conglomerados analizar, se hace uso del coeficiente de aglomeración y del dendograma.

Se analiza el resultado de la clasificación usando Anova, técnicas de comparaciones múltiples, gráficos de medias, gráfico de individuos en conglomerados y las tablas de contingencia. Por último, se valida la solución de clasificación con el método no jerárquico umbral paralelo y con variables no incluidas en el modelo.

* Tesis de grado

** Facultad de Ciencias. Escuela de Matemáticas. Especialización en Estadística. Director: Germán Moreno Arenas. Doctor en Estadística

SUMMARY

Title: Patterns of employment of the graduates of 2008 from the Pontificia Bolivariana University, from one year to two years after graduating: Analysis of classification*

Author: Diana Liceth Garcés Portilla**

Keywords: Graduates, employment characteristics, factor analysis and classification.

This current research pretends to gather graduates of 2008 from the Pontificia Bolivariana University, from one year to two years after graduating, in groups with similar and unlike job characteristics.

This reduction of graduates in groups facilitates the interpretation of data, as patterns of employment are established. To get to the definition of profiles, several stages were carried out, including, the reduction of correlated variables through factor analysis to enter them into the classification analysis. After that, the similarity measure was determined, which is related to the type of variable of study. In this case, the Euclidean distance squared is used, and some variables are quantitative and others, because of the escalation carried out by using Likert scale, may be treated as metric.

As soon as the similarity measure of distance was determined, the method of Ward conglomerates formation was applied, as it is a procedure that is based on the Euclidean distance measure squared. When established the structure of conglomerates, in order to define how many conglomerates are to be analyzed, a coefficient of agglomeration and a dendrogram are used.

The result of the classification is analyzed by using ANOVA, multiple comparison techniques, measure charts, and contingency tables. Lastly, the solution of classification is validated with a parallel non-hierarchical threshold method with variables that are not included in the modeling.

* Thesis

** Faculty School of Mathematics. Specialization in Statistics. Directed by German Moreno Arenas. Ph.D. in Statistics

1. INTRODUCCIÓN

A finales de la década de los noventa del siglo XX, los ministerios de educación de la Unión Europea se preocuparon por evaluar si sus sistemas educativos universitarios estaban formando profesionales con las competencias suficientes para aportar soluciones a los problemas de los sectores productivos y de la sociedad. Este hecho motivó investigaciones relacionadas con la transición del recién graduado al mercado laboral, contribuyendo al conocimiento de los procesos de inserción, la situación en el empleo, el tipo de competencias requeridas en el desempeño profesional y el impacto de la formación recibida durante su permanencia en la universidad.

Como una consecuencia de este interés, se desarrollaron proyectos en diversos países europeos orientados a conocer la situación de los graduados de instituciones de educación superior: el primero, el proyecto CHEERS¹ (1997, financiado por el IV Programa Marco de la UE²), el proyecto Italiano ALMALAUREA³ (1998-2011), el proyecto REFLEX⁴ (2004, VI Programa Marco de la UE) en el cual participaron graduados de quince países Europeos y Japón, recopilando información mediante una encuesta a graduados con cinco años de haber recibido el título profesional y la Red GRADUA2⁵ (2006). Estos mismos estudios se replicaron para Latinoamérica a través del proyecto PROFLEX⁶ (2006, Programa de Cooperación entre IES⁷ de la Unión Europea y América Latina).

Tomando como referencia estos estudios, en el 2008 el Ministerio de Educación Nacional de Colombia, declaró a la pertinencia como uno de los pilares esenciales

¹ Career After Higher Education: A European Research Study.

² Unión Europea.

³ Es un servicio para los graduados, fundado por el Observatorio Estadístico de la Universidad de Bolonia para poner en la web las hojas de vida de los graduados de todas las universidades Italianas y hacer los estudios de graduados.

⁴ The Flexible Professional in the Knowledge Society.

⁵ Red de Universidades Europeas y Latinoamericanas para hacer seguimiento a sus graduados.

⁶ Profesional Flexible en la Sociedad del Conocimiento.

⁷ Instituciones de Educación Superior.

para mejorar la calidad de educación universitaria. Por esta razón se crea el Observatorio Laboral para la Educación y en común acuerdo con las instituciones de educación superior se diseñaron encuestas de seguimiento a graduados para ser aplicadas en cuatro momentos: en el momento del grado, un año, tres y cinco años después de recibir el título profesional.

La Universidad Pontificia Bolivariana de Bucaramanga (UPB) en correspondencia con los lineamientos del Ministerio de Educación Nacional, tomó en el 2008 la decisión de realizar estudios de seguimiento a graduados como elemento principal de la evaluación de la calidad de la formación impartida.

En general, los estudios existentes a nivel nacional e internacional solo han presentado resultados descriptivos univariados y bivariados, la intención de este proyecto es presentar un análisis multivariado, con el propósito de establecer los perfiles del empleo de los estudiantes de la UPB entre un año y dos años después del grado, tomando como base los datos recopilados de los egresados de las cohortes 2008 generados de la aplicación de una encuesta de seguimiento en los meses de diciembre de 2009 a enero de 2010.

La población de estudio de dichas cohortes fue de 465 graduados, la muestra obtenida para el estudio de seguimiento un año después del grado fue de 130 graduados. De esta muestra el 69.2% (89) de los graduados se encontraban empleados⁸ en empresas privadas o públicas, los cuales son el objeto de estudio del presente análisis de clasificación.

Las variables determinantes para el análisis de la situación del empleo de los graduados de los estudios internacionales, los de Ministerio de Educación Nacional de Colombia y los de la UPB Bucaramanga, son comparables en la

⁸ Entendiéndose empleado como “Persona con un vínculo laboral con el empleador de manifiesto en un contrato de trabajo verbal o escrito.

mayoría de variables, teniendo en cuenta que las experiencias de seguimiento internacionales sirvieron de marco de referencia para los estudios colombianos. Las variables relacionadas con la situación laboral de los graduados tomadas como referencia para la determinación de los perfiles son: el salario bruto mensual en el trabajo actual, relación del trabajo con el programa de educación superior culminado, el grado de utilidad en el trabajo de los conocimientos, habilidades y destrezas aprendidas en el programa de educación superior en que se graduó, el grado de satisfacción con la ocupación actual, la diferencia entre las expectativas de un buen empleo según el graduado y las características reales del trabajo actual desde el punto de vista de la estabilidad laboral, la mejora de los ingresos y las perspectivas de crecimiento profesional (oportunidades de adquirir nuevos conocimientos y tener la posibilidad de afrontar nuevos retos) (Tabla 1).

Tabla 1. Variables relacionadas con la situación laboral

ID	PREGUNTA DE INVESTIGACIÓN	VARIABLE	TIPO DE VARIABLE	TIPO DE ESCALA	ESCALAMIENTO																																								
A1	¿Cuál fue su ingreso laboral el mes pasado? (No incluya horas extras, viáticos, ni ayudas en especie)	Salario bruto mensual ⁹ .	Cuantitativa	Razón																																									
A2	¿Qué tan relacionado es su trabajo con el programa que terminó?	Relación del trabajo con el programa de educación superior culminado.	Cualitativa	Ordinal	1. <input type="checkbox"/> Nada relacionado 2. <input type="checkbox"/> Poco relacionado 3. <input type="checkbox"/> Medianamente relacionado 4. <input type="checkbox"/> Muy relacionado																																								
A3	¿Qué tan útiles han sido en su trabajo los conocimientos, habilidades y destrezas aprendidas en el programa de educación en que se graduó?	Utilidad del programa de educación superior en el trabajo.	Cualitativa	Ordinal	1. <input type="checkbox"/> Nada útil 2. <input type="checkbox"/> Poco útil 3. <input type="checkbox"/> Medianamente útil 4. <input type="checkbox"/> Muy útil																																								
A4	¿Cómo calificaría su satisfacción con el trabajo actual?	Satisfacción con el trabajo.	Cualitativa	Ordinal	<table border="1"> <tr> <td>1. Muy insatisfecho</td> <td>2. Insatisfecho</td> <td>3. Medianamente satisfecho</td> <td>4. Satisfecho</td> <td>5. Muy satisfecho</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table>	1. Muy insatisfecho	2. Insatisfecho	3. Medianamente satisfecho	4. Satisfecho	5. Muy satisfecho	1	2	3	4	5	<input type="checkbox"/>																													
1. Muy insatisfecho	2. Insatisfecho	3. Medianamente satisfecho	4. Satisfecho	5. Muy satisfecho																																									
1	2	3	4	5																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																									
A5	En su opinión, en este momento que tanta importancia tiene para usted y para su empresa la estabilidad laboral.	Diferencia entre Importancia dada por el graduado y la empresa a la estabilidad laboral.	Cuantitativa	Intervalo	<table border="1"> <tr> <td colspan="5">1. Nivel de importancia para Ud.</td> </tr> <tr> <td>Nada importante</td> <td colspan="3"></td> <td>Muy importante</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td colspan="5">2. Nivel de importancia para la empresa.</td> </tr> <tr> <td>Nada importante</td> <td colspan="3"></td> <td>Muy importante</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table>	1. Nivel de importancia para Ud.					Nada importante				Muy importante	1	2	3	4	5	<input type="checkbox"/>	2. Nivel de importancia para la empresa.					Nada importante				Muy importante	1	2	3	4	5	<input type="checkbox"/>								
1. Nivel de importancia para Ud.																																													
Nada importante				Muy importante																																									
1	2	3	4	5																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																									
2. Nivel de importancia para la empresa.																																													
Nada importante				Muy importante																																									
1	2	3	4	5																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																									

⁹ Para hacer comparables los salarios de los graduados que están en el exterior y los radicados en Colombia, se utilizó la paridad del poder adquisitivo (PPA).

ID	PREGUNTA DE INVESTIGACIÓN	VARIABLE	TIPO DE VARIABLE	TIPO DE ESCALA	ESCALAMIENTO																																								
A6	En su opinión, en este momento que tanta importancia tiene para usted y para su empresa la oportunidad de adquirir nuevos conocimientos.	Diferencia entre la importancia dada por el graduado y la empresa a la oportunidad de adquirir nuevos conocimientos.	Cuantitativa	Intervalo	<table border="1"> <tr> <td colspan="5">1. Nivel de importancia para Ud.</td> </tr> <tr> <td>Nada importante</td> <td colspan="3"></td> <td>Muy importante</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table> <table border="1"> <tr> <td colspan="5">2. Nivel de importancia para la empresa.</td> </tr> <tr> <td>Nada importante</td> <td colspan="3"></td> <td>Muy importante</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table>	1. Nivel de importancia para Ud.					Nada importante				Muy importante	1	2	3	4	5	<input type="checkbox"/>	2. Nivel de importancia para la empresa.					Nada importante				Muy importante	1	2	3	4	5	<input type="checkbox"/>								
1. Nivel de importancia para Ud.																																													
Nada importante				Muy importante																																									
1	2	3	4	5																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																									
2. Nivel de importancia para la empresa.																																													
Nada importante				Muy importante																																									
1	2	3	4	5																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																									
A7	En su opinión, en este momento que tanta importancia tiene para usted y para su empresa la mejora de los ingresos laborales.	Diferencia entre la importancia dada por el graduado y la empresa a la mejora de los ingresos laborales.	Cuantitativa	Intervalo	<table border="1"> <tr> <td colspan="5">1. Nivel de importancia para Ud.</td> </tr> <tr> <td>Nada importante</td> <td colspan="3"></td> <td>Muy importante</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table> <table border="1"> <tr> <td colspan="5">2. Nivel de importancia para la empresa.</td> </tr> <tr> <td>Nada importante</td> <td colspan="3"></td> <td>Muy importante</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table>	1. Nivel de importancia para Ud.					Nada importante				Muy importante	1	2	3	4	5	<input type="checkbox"/>	2. Nivel de importancia para la empresa.					Nada importante				Muy importante	1	2	3	4	5	<input type="checkbox"/>								
1. Nivel de importancia para Ud.																																													
Nada importante				Muy importante																																									
1	2	3	4	5																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																									
2. Nivel de importancia para la empresa.																																													
Nada importante				Muy importante																																									
1	2	3	4	5																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																									
A8	En su opinión, en este momento que tanta importancia tiene para usted y para su empresa la posibilidad de afrontar nuevos retos.	Diferencia entre la importancia dada por el graduado y la empresa a la posibilidad de afrontar nuevos retos.	Cuantitativa	Intervalo	<table border="1"> <tr> <td colspan="5">1. Nivel de importancia para Ud.</td> </tr> <tr> <td>Nada importante</td> <td colspan="3"></td> <td>Muy importante</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table> <table border="1"> <tr> <td colspan="5">2. Nivel de importancia para la empresa.</td> </tr> <tr> <td>Nada importante</td> <td colspan="3"></td> <td>Muy importante</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table>	1. Nivel de importancia para Ud.					Nada importante				Muy importante	1	2	3	4	5	<input type="checkbox"/>	2. Nivel de importancia para la empresa.					Nada importante				Muy importante	1	2	3	4	5	<input type="checkbox"/>								
1. Nivel de importancia para Ud.																																													
Nada importante				Muy importante																																									
1	2	3	4	5																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																									
2. Nivel de importancia para la empresa.																																													
Nada importante				Muy importante																																									
1	2	3	4	5																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																									

El método estadístico utilizado para la determinación de los perfiles del empleo de los graduados de la UPB entre un año y dos después del grado, es el análisis de clasificación, previa aplicación del análisis factorial.

Una vez definidos los perfiles con las variables anteriormente mencionadas, se determinará si hay diferencias significativas entre ellos con relación a las características, género, escuela y facultad y las variables de ocupación laboral: tipo de contrato, nivel del cargo actual y tamaño de la empresa.

El software estadístico de apoyo para la aplicación de la técnica multivariante es el SPSS 17.

En el numeral 2, se presenta el marco de referencia para el análisis multivariado factorial y de Clasificación.

En el numeral 3, se realiza el análisis factorial y de clasificación. Por último, se incluyen las conclusiones, recomendaciones y anexos.

2. MARCO DE REFERENCIA

El análisis de clasificación, busca agrupar o formar conjuntos de objetos con características similares y a la vez que los diferentes grupos formados sean los más disímiles posibles entre sí. Así, el análisis de clasificación tiene como objetivo principal, definir la estructura de los datos colocando las observaciones más parecidas en grupos.

Para poder lograr este objetivo se deben considerar los siguientes aspectos:

2.1 SUPUESTOS DEL ANÁLISIS DE CLASIFICACIÓN

2.1.1 Representatividad de la muestra. La muestra obtenida debe ser representativa de la población, teniendo en cuenta que el análisis de clasificación es tan bueno como la representatividad de la muestra (Hair, 2004).

2.1.2 Multicolinealidad. El efecto que tiene las variables multicolineales sobre el análisis de clasificación, es que en esta técnica multivariante están implícitamente ponderadas con más fuerza y lo que se busca es reflejar el efecto de cada una de las variables sobre la formación de la estructura de los conglomerados de manera independiente (Hair, 2004).

Una de las maneras de detectar la multicolinealidad en las variables es el uso del factor de tolerancia y su inverso, el factor de inflación de la varianza (VIF). Estas medidas dan el grado en el que cada variable independiente se explica por otras variables independientes. En términos sencillos, cada variable independiente se convierte en una variable dependiente y se realiza la regresión con el resto de las variables independientes.

La tolerancia es la cantidad de variabilidad de las variables independientes seleccionadas no explicada por el resto de las variables independientes. Por lo tanto, un valor reducido (y elevados valores de VIF) denotan una alta colinealidad. Un límite de tolerancia sugerido es 0,10 (o por encima de un VIF de 10) corresponde a una correlación múltiple de 0,95 donde la mayor correlación es uno. Cuando se desea evaluar la asociación lineal de parejas de variables o bivariadas, se aplica la correlación de Pearson, sólo si las variables son métricas y cumplen con el supuesto de distribución normal, de lo contrario, se aplica la prueba no paramétrica; la correlación de Spearman. Al aplicar la correlación de Pearson sin cumplir con el supuesto de distribución normal y que las variables sean métricas, el resultado de la correlación es una subestimación de verdadero valor del parámetro de correlación.

Si se detecta multicolinealidad una manera de solucionar el problema es reducir las variables originales en un conjunto reducido de variables compuestas (factores) que no se correlacionan entre sí, es decir, aplicar un análisis factorial. Hay diferentes formas de verificar si el modelo factorial es apropiado. Una de ellas es el examen visual de la matriz de correlación. Un número sustancial de correlaciones mayores que 0.30, implica que el modelo es adecuado.

Además de la inspección visual existen estadísticas formales, como la prueba de esfericidad de Bartlett. Esta comprueba la hipótesis nula de que las variables no se correlacionan en la población. Un valor grande de la estadística de prueba, favorecerá el rechazo de la hipótesis nula. Si no se puede rechazar la hipótesis nula, hay dudas de que el análisis factorial sea apropiado.

Otra estadística, es la medida de adecuación de la muestra de Kaiser-Meyer-Olkin (KMO), valores pequeños de KMO indican que las correlaciones entre pares de variables no se explican mediante otras variables y el análisis factorial posiblemente no sea apropiado. Es deseable un valor de KMO mayor de 0.5.

Por último, la medida de adecuación muestral (MSA). Este índice se extiende de 0 a 1, llegando a uno cuando la variable es perfectamente predicha sin error por las otras variables. La medida puede ser interpretada con las siguientes directrices: 0.80 o superior, sobresaliente; 0.70 o superior, regular; 0.60 o superior, mediocre; 0.50 o superior, despreciable; y por debajo de 0,50 inaceptable.

Una vez definido si el modelo factorial es adecuado, se siguen las siguientes etapas en el análisis multivariado factorial:

Determinación del método de análisis factorial

Una vez que se ha determinado que el análisis factorial es adecuado, debe elegirse el método apropiado de análisis factorial. Los métodos del análisis factorial se distinguen por la manera de derivar los pesos o los coeficientes de las puntuaciones de los factores. Los dos métodos básicos son el análisis de componentes principales y el análisis de los factores comunes.

En el análisis de los componentes principales se considera la varianza total de los datos. La diagonal de la matriz de correlación consta de unidades y en la matriz de los factores se incluye toda la varianza.

El análisis de componentes principales se recomienda cuando lo que más interesa es determinar el número mínimo de factores que explicarán la mayor varianza de los datos para uso en el análisis multivariado siguiente. Análisis que se aplica para este estudio.

Determinación del número de factores

Uno de los procedimientos para decidir sobre el número de factores del modelo factorial, es la determinación basada en el porcentaje de varianza. En este método, se determina el número de factores extraídos de modo que el porcentaje acumulado de la varianza extraída por los factores llegue a cierto nivel

satisfactorio. Cuál sea el nivel satisfactorio de varianza depende del problema; sin embargo, se recomienda que los factores extraídos expliquen por lo menos el 60% de la varianza.

Rotación de Factores

Aunque la matriz factorial inicial sin rotación, indica la relación o la correlación entre las variables individuales y los factores (Cargas factoriales), casi nunca da factores que puedan interpretarse, porque estos se correlacionan con muchas variables. Al rotar los factores, se distribuyen las cargas factoriales en los factores y es más fácil la lectura de estos, porque cada variable tiene una carga factorial sobre uno o pocos factores diferentes de cero.

El método de rotación más usado es el procedimiento Varimax, un método ortogonal de rotación que reduce al mínimo el número de variables con cargas grandes en un factor, lo que mejora la capacidad de interpretación de los factores.

Interpretación de los factores

Una carga de una variable sobre un factor es significativa, dependiendo del tamaño muestral, el nivel de significancia establecido, la potencia y el error estándar de las cargas factoriales. Existe una tabla, donde se dan las directrices para la identificación de las cargas factoriales significativas basadas en el tamaño muestral, con nivel de significación del 0,05 y un nivel de potencia del 80% (Anexo B).

Para la muestra de 89 datos del presente estudio, para que una carga factorial sea significativa debe ser del 0.589.

Si un factor no puede definirse claramente con las variables originales, debe designarse como valor indefinido o general.

Cuando se ha obtenido una solución factorial en que todas las variables tienen una carga significativa sobre un factor, el investigador procura atribuir un significado al patrón de cargas factoriales, es decir, una etiqueta que represente al factor, en referencia a las variables que tengan mayores cargas factoriales.

Los signos de las cargas factoriales en cada factor se interpretan como otros coeficientes de correlación, signos iguales significa que las variables están positivamente relacionadas, signos opuestos significan que las variables están negativamente relacionadas.

Cálculo de la puntuación de los factores

Si la finalidad es reducir las variables originales a un conjunto menor de variables compuestas (factores) para el siguiente análisis multivariado, es útil calcular las puntuaciones de los factores de cada encuestado. Un factor no es más que una combinación lineal de las variables originales.

Determinación del ajuste del modelo

La última etapa del análisis factorial es la determinación del ajuste del modelo. Una premisa básica del análisis factorial es que la correlación observada entre las variables puede atribuirse a factores. Por consiguiente, las correlaciones entre las variables se deducen o reproducen de las correlaciones estimadas entre las variables y los factores. Las diferencias entre las correlaciones observadas (Matriz de correlación) y las correlaciones reproducidas (que se estiman en la matriz factorial) se examinan para determinar el ajuste del modelo. Estas diferencias se llaman residuos.

Si hay muchos residuos mayores de 0,05, el modelo de los factores no da un buen ajuste para los datos y hay que volver a considerarlo.

2.2 MEDIDA DE SIMILARIDAD

Esta medida permite establecer el parecido o la proximidad entre objetos que van a ser agrupados, en términos de las variables que los describen.

La medida de similitud depende del tipo de escala de medición de las variables que definen las características de los objetos. Cuando las variables están medidas en escalas de razón o de intervalos, es decir métricas (cuantitativas), se utilizan medidas de correlación y distancia.

Las medidas de correlación representan la similitud mediante la correspondencia de patrones entre las características y no observa las magnitudes. Por lo tanto, las medidas de correlación se utilizan rara vez porque el interés de la mayoría de las aplicaciones del análisis de clasificación está en las magnitudes de los objetos, no en los patrones de los valores (Hair, 2004).

La medida de distancia más utilizada es la distancia euclidiana. Las medidas de distancia son en realidad medidas de diferencia, donde los valores pequeños indican mayor similitud. Otra de las medidas de distancia particularmente usadas es la distancia euclidiana al cuadrado, que es la suma de las diferencias entre las variables que describen los objetos al cuadrado.

Un problema al que se enfrenta todas las medidas de distancia es que el uso de datos no estandarizados implica inconsistencias entre las soluciones de clasificación cuando cambia la escala de las variables.

Las medidas de asociación o de similitud se utilizan para comparar objetos cuyas características están medidas en escalas nominales y ordinales, es decir variables no métricas (cualitativas).

La medida de similitud más usada en el caso de datos nominales dicotómicos, es el coeficiente de asociación de Jaccard, que describe la razón entre el número de presencias de los atributos en los dos objetos, dividido por la sumatoria del número de presencias de los atributos en los dos objetos, más el número de veces en que los atributos fueron ausentes en un objeto y presentes en el otro (y su inversa).

Los coeficientes de asociación tienen el limitante que sólo puede aplicarse a respuestas dicotómicas. Las variables cualitativas con varios niveles de respuesta, es posible tratarlas de manera similar a los datos dicotómicos, con cada nivel de la variable siendo considerada como una variable dicotómica (Everitt, 1974).

La medida de similitud en el caso de variables cualitativas con varios niveles de respuesta suelen estar basadas en la χ^2 (Chi-Cuadrado) de Pearson.

El coeficiente de similitud de Gower (1971) es una medida de similitud que permite la utilización simultánea de variables cuantitativas y cualitativas. Aplicando este coeficiente de similitud se puede determinar el grado de semejanza o similitud entre individuos; a los cuales se le han medido características cuantitativas, cualitativas binarias o de varias categorías.

Para efectos del presente estudio se utilizará la distancia euclidiana al cuadrado, aunque no todas las variables son de naturaleza cuantitativa (métricas), como es el caso de las variables ordinales: A2 (Relación del trabajo con el programa de educación superior culminado), A3 (Utilidad del programa de educación superior en el trabajo) y A4 (Satisfacción con el trabajo). Pero su escalamiento, en forma de escala de Likert, permite tratarlas como variables métricas, para estudios de análisis de clasificación (conglomerados) y análisis factorial (Malhotra, 2008).

2.3 LA FORMACIÓN DE LOS CONGLOMERADOS

El método de formación de conglomerados, es un algoritmo de aglomeración mediante el cual se define la estructura de los datos, es decir se agrupan los objetos similares en un conglomerado específico.

Hay diferentes procedimientos de formación de conglomerados o algoritmos, pero el criterio esencial, es que intentan maximizar las diferencias entre los conglomerados relativa a la variación dentro de los conglomerados (Hair, 2004).

Los procedimientos de conglomerado pueden ser jerárquicos y no jerárquicos. Los métodos jerárquicos consisten en la construcción de una estructura en forma de árbol. Esta estructura puede formarse por un proceso aglomerativo o por división.

En el proceso aglomerativo cada uno de los objetos es un conglomerado (grupo unitario), luego los grupos más cercanos se unen para formar un solo conglomerado, así sucesivamente hasta que el último conglomerado está formado por todos los objetos.

Los métodos por división¹⁰, empiezan con un solo conglomerado formado por todos los objetos, luego se dividen en dos grupos, estos en otros dos, hasta que cada objeto llega a ser un conglomerado.

Los métodos aglomerativos más habituales son: Enlace simple, enlace completo, enlace promedio, método de Ward y del centroide.

El enlace simple se basa en la distancia mínima. Empieza con los dos objetos que tengan la mínima distancia entre sí o dentro de un límite de similaridad dispuesto.

¹⁰ Dado que los programas informáticos más habituales utilizan los procedimientos aglomerativos, y los métodos por división actúan como métodos aglomerativos al revés, se centra la explicación en los procedimientos aglomerativos.

Luego puede ocurrir que la mínima distancia sea entre el grupo ya formado y otro objeto o entre dos objetos cercanos. Se sigue así sucesivamente hasta que todos los objetos hacen parte de un conglomerado. El uso de la distancia mínima puede unir conglomerados que no son homogéneos entre sí.

El enlace completo se basa en la distancia máxima. En este caso dos grupos son mezclados solo si sus objetos más distantes están lo suficiente cerca de manera conjunta.

El enlace promedio, comienza igual que el enlace simple, pero el criterio de aglomeración es la distancia promedio de todos los objetos de un conglomerado con todos los objetos del otro. Este método no depende de valores extremos, como en el caso del enlace simple y completo. Además la partición se basa en todos los miembros de los conglomerados en lugar de un par único de miembros extremos.

En el Método Ward, se calculan las medias de todas las variables de cada conglomerado. Luego se calcula la distancia cuadrada euclidiana de cada objeto a la media del conglomerado, estas distancias se suman para todos los objetos. En cada etapa se combinan los dos conglomerados con el menor incremento en la suma general de cuadrados de las distancias (Malhotra, 2008). Se define matemáticamente como:

$$SCW = \frac{1}{1/n_h + 1/n_k} \| \bar{X}_h - \bar{X}_k \|^2$$

donde \bar{X}_h y \bar{X}_k son los centroides, n_h y n_k son los tamaños de los conglomerados h y k respectivamente.

En el método del centroide, la distancia entre dos conglomerados es la distancia entre sus centroides (que son las medias de todas las variables). Cada vez que se agrupan los objetos se calcula un nuevo centroide.

De los métodos jerárquicos, los de enlace promedio y Ward han resultado mejores que los demás (Malhotra, 2008). Además no se tan sensibles a los datos atípicos.

Los métodos no jerárquicos, no implican la formación de árboles. Se empieza con la definición del número de conglomerados a formar.

Los métodos no jerárquicos son: Umbral secuencial, umbral paralelo y optimización. El umbral secuencial, empieza con un punto semilla (centro) de conglomerado que incluye todos los objetos que caen dentro de una distancia previamente especificada. Cuando ya todos objetos están incluidos en esa distancia, se selecciona una segunda semilla de conglomerado y se sigue de la misma manera como el procedimiento descrito para la primera semilla. Cuando un objeto esta conglomerado en una semilla, ya no se considera para un conglomerado de otras semillas.

El en el umbral paralelo, se seleccionan varias semillas simultáneamente desde el principio y asigna objetos dentro de la distancia umbral hasta la semilla más cercana. A medida que avanza el proceso, se pueden ajustar las distancias umbral para incluir más o menos objetos en los conglomerados.

El procedimiento de optimización difiere de los dos métodos del umbral en que los objetos pueden volver a asignarse a otros conglomerados.

La desventaja de los métodos no jerárquicos es que deben especificarse el número de conglomerados desde el principio y la selección de los puntos semilla

(centros) es arbitraria, es decir, los resultados del conglomerado dependen de cómo se elijan los centros.

Cada metodología de formación de conglomerados aplicada al mismo conjunto de datos generalmente arroja resultados diferentes. Para obviar parte de esta dificultad, se debe emplear un procedimiento concordante con la naturaleza de la tipología esperada, con las variables a considerar y la medida de similaridad usada (Díaz Monroy, 2000).

Para el presente estudio, se utiliza el método Ward, teniendo en cuenta que la medida de similitud de distancia seleccionada es la distancia euclidiana al cuadrado.

2.4 NÚMERO DE CONGLOMERADOS

El coeficiente de aglomeración o el cambio porcentual del coeficiente de aglomeración de un nivel a otro o de un conglomerado a otro, es particularmente utilizado para determinar el número de conglomerados. Los coeficientes pequeños indican que están surgiendo conglomerados claramente homogéneos. Unir dos conglomerados muy diferentes produce un coeficiente muy grande o un gran porcentaje de cambio en el coeficiente.

2.5 INTERPRETACIÓN DE LOS CONGLOMERADOS

Se tomará la decisión del número de conglomerados finales, determinando si hay diferencias significativas entre los conglomerados respecto a las características que los definen, por medio de una Anova. El número de conglomerados que mejor agrupe a los graduados, es decir, hayan mayores diferencias significativas entre los conglomerados, será la solución final de clasificación.

Con la solución final de clasificación, para definir el perfil de cada conglomerado, se describen mediante el análisis de la gráfica de perfil y las tablas de contingencia de los conglomerados y las variables.

2.6 LA VALIDEZ DE LA SOLUCIÓN DE CLASIFICACIÓN

Aunque no existe un método único para asegurar la validez y la relevancia práctica de la solución final de conglomerado, diferentes autores han propuesto formas de evaluar la solución final. A continuación se plantea la del autor Joseph Hair (Hair, 2004).

Primero, se evalúa la solución mediante la aplicación de métodos de formación de conglomerados alternativos y comparando las soluciones. A continuación se evalúan los conglomerados sobre medidas adicionales, es decir, con variables no incluidas en el modelo.

En el primer paso, para evaluar la validez de la solución de clasificación en conglomerados, se utiliza el método no jerárquico denominado umbral paralelo, donde el número conglomerados corresponde a los determinados en el método Ward y los puntos semilla (centros) son los centroides de cada uno de los conglomerados obtenidos de la solución por el método Ward.

En el segundo paso, para evaluar la validez predictiva de los conglomerados, se basa, en la detección de diferencias significativas entre los conglomerados y las variables no incluidas en el análisis, pero que guardan relación con el estudio, y que pueden ser importantes en la diferenciación de los conglomerados como las variables demográficas.

Para evaluar la diferencia entre las variables categóricas adicionales y los conglomerados, se utiliza la prueba χ^2 (Chi-Cuadrado) y si las diferencias son

significativas se aplica el coeficiente V de Cramer para establecer la fuerza de asociación entre las variables. Este coeficiente se usa para tablas de contingencia de cualquier tamaño. La V de Cramer va de 0 a 1, donde un valor cercano a uno, indica un alto grado de asociación.

3. RESULTADOS

3.1 SUPUESTOS DEL ANÁLISIS DE CLASIFICACIÓN

3.1.1 Representatividad de la muestra. La población de estudio estuvo conformada por 465 graduados que recibieron el grado en el 2008. Se tomó una muestra de 130 graduados con un error de estimación del 5.2%, un nivel de confianza del 90% y se calculó un $p= 0,7$ y $q= 0.3$ con base a una premuestra para la variable situación laboral del graduado, encontrándose que el 70% tenían un empleo en el momento de realizar el estudio. De la muestra seleccionada se tomó el 69.2% (89) de los graduados se encontraban empleados en empresas privadas o públicas, los cuales son el objeto de estudio del presente análisis de clasificación.

3.1.2 Multicolinealidad. Se observa en la matriz de correlación, la presencia de nueve asociaciones significativas entre las veintiocho correlaciones, es decir, el 32.14%, tienen un coeficiente de Spearman mayor que 0.30 (Tabla 2. Subrayadas en rojo) a un nivel de significancia de 0.05. Lo cual indica que hay correlación entre las variables.

Tabla 2. Matriz de correlación

			Correlaciones							
			A1	A2	A3	A4	A5	A6	A7	A8
Rho de Spearman	A1	Coefficiente de correlación	1.000	.178	.040	.390**	.093	-.119	-.181	-.069
		Sig. (bilateral)	.	.095	.712	.000	.384	.268	.089	.519
		N	89	89	89	89	89	89	89	89
	A2	Coefficiente de correlación	.178	1.000	.357**	.403**	.104	-.128	-.213*	-.036
		Sig. (bilateral)	.095	.	.001	.000	.333	.231	.045	.737
		N	89	89	89	89	89	89	89	89
	A3	Coefficiente de correlación	.040	.357**	1.000	.221*	-.051	-.088	-.049	.012
		Sig. (bilateral)	.712	.001	.	.037	.636	.413	.651	.910
N		89	89	89	89	89	89	89	89	
A4	Coefficiente de correlación	.390**	.403**	.221*	1.000	-.014	-.270*	-.249*	-.147	
	Sig. (bilateral)	.000	.000	.037	.	.896	.010	.019	.168	
	N	89	89	89	89	89	89	89	89	
A5	Coefficiente de correlación	.093	.104	-.051	-.014	1.000	.301**	.437**	.310**	
	Sig. (bilateral)	.384	.333	.636	.896	.	.004	.000	.003	
	N	89	89	89	89	89	89	89	89	
A6	Coefficiente de correlación	-.119	-.128	-.088	-.270*	.301**	1.000	.422**	.604**	
	Sig. (bilateral)	.268	.231	.413	.010	.004	.	.000	.000	
	N	89	89	89	89	89	89	89	89	
A7	Coefficiente de correlación	-.181	-.213*	-.049	-.249*	.437**	.422**	1.000	.406**	
	Sig. (bilateral)	.089	.045	.651	.019	.000	.000	.	.000	
	N	89	89	89	89	89	89	89	89	
A8	Coefficiente de correlación	-.069	-.036	.012	-.147	.310**	.604**	.406**	1.000	
	Sig. (bilateral)	.519	.737	.910	.168	.003	.000	.000	.	
	N	89	89	89	89	89	89	89	89	

** . La correlación es significativa al nivel 0,01 (bilateral).

* . La correlación es significativa al nivel 0,05 (bilateral).

Teniendo en cuenta que la multicolinealidad se halla, tomando cada variable independiente como una variable dependiente y realizando la regresión con el resto de las variables independientes. Para aplicar la regresión debe cumplirse con el supuesto de normalidad de las variables, lo cual no se cumple en este caso (Anexo A), por lo tanto no se hace lectura del factor de tolerancia y su inverso, el factor de inflación de la varianza (VIF).

Se toma como prueba suficiente las correlaciones bivariadas, para eliminar su efecto sobre el análisis de clasificación, utilizando las variables reducidas del análisis factorial.

Además se apoya el uso de variables reducidas con la aplicación de las estadísticas formales; la prueba de esfericidad de Bartlett, la medida de adecuación muestral KMO y la medida de adecuación muestral MSA.

La prueba de esfericidad es significativa (Tabla 3), por lo tanto, las variables están correlacionadas en la población. Además la prueba KMO (0.653) es mayor que 0.5, lo que significa, que las correlaciones entre pares de variables se explican mediante otras variables (Tabla 3).

Tabla 3. KMO y Prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		.653
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	130.047
	gl	28
	Sig.	.000

Como complemento, la medida de adecuación muestral (MSA), es aceptable para cada una de las variables, es decir, son superiores a 0,5¹¹ (Tabla 4). Por lo tanto el modelo factorial es adecuado.

Tabla 4. Matriz de correlación anti-imagen

	A1	A2	A3	A4	A5	A6	A7	A8
A1	.583 ^a	,071	-,001	-,335	-,135	,069	,129	-,065
A2	,071	.539 ^a	-,324	-,288	-,157	-,016	,159	-,066
A3	-,001	-,324	.532 ^a	-,052	,078	,050	-,110	-,022
A4	-,335	-,288	-,052	.658 ^a	-,082	,127	,130	,022
A5	-,135	-,157	,078	-,082	.641 ^a	-,122	-,337	-,049
A6	,069	-,016	,050	,127	-,122	.700 ^a	-,127	-,496
A7	,129	,159	-,110	,130	-,337	-,127	.712 ^a	-,241
A8	-,065	-,066	-,022	,022	-,049	-,496	-,241	.678 ^a

a. Medida de adecuación muestral

El método de análisis factorial que se utiliza en el presente estudio, es el análisis de componentes principales, ya que el objetivo es, utilizar unas variables finales que son una combinación lineal de las variables originales, para tratar el problema de multicolinealidad y poder aplicar el análisis de clasificación¹².

¹¹ HAIR Joseph F; ANDERSON Rolph E, Análisis Multivariante, Quinta edición, Madrid: Prentice Hall, 2004, pág. 88 y 108.

¹² MALHOTRA Naresh K, Investigación de Mercados, Quinta edición, Mexico: Prentice Hall, 2008, pág. 564.

En la tabla 5 de la varianza total explicada por los componentes principales, se observa que los tres primeros componentes, explican el 64.559%. Valor que se encuentra por encima de lo recomendado, por ende los tres primeros componentes dan una buena explicación de la variabilidad total y se tendrán en cuenta en la interpretación de las correlaciones de las variables con cada uno de los componentes principales (Cargas Factoriales).

Tabla 5. Varianza total explicada por los componentes principales

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2.392	29.895	29.895	2.392	29.895	29.895
2	1.675	20.943	50.838	1.675	20.943	50.838
3	1.098	13.721	64.559	1.098	13.721	64.559
4	.774	9.669	74.229			
5	.715	8.932	83.161			
6	.522	6.521	89.682			
7	.447	5.584	95.265			
8	.379	4.735	100.000			

Método de extracción: Análisis de Componentes principales.

La matriz de componentes principales no rotada (Tabla 6), no permite la interpretación de los factores, ya que la varianza extraída se distribuye igual o de manera similar en dos o más componentes. Es el caso de la variable A1 (Salario bruto mensual) y A3 (Utilidad del programa de educación superior en el trabajo), se distribuyen de manera similar en el componente 2 y 3, Además A4 (Satisfacción con el trabajo) y A5 (Diferencia entre la importancia dada por el graduado y la empresa a la estabilidad laboral), se puede afirmar que tienen una carga factorial igual en el componente 1 y 2.

En cambio la matriz de componentes rotada (Tabla 7), permite identificar que las variables A5 (Diferencia entre la importancia dada por el graduado y la empresa a la estabilidad laboral), A6 (Diferencia entre la importancia dada por el graduado y la empresa a la oportunidad de adquirir nuevos conocimientos), A7 (Diferencia entre la importancia dada por el graduado y la empresa a la mejorar de los

ingresos laborales), A8 (Diferencia entre la importancia dada por el graduado y la empresa a la posibilidad de afrontar nuevos retos) se relacionan más fuerte con el componente 1, A4 (Satisfacción con el trabajo) y A1 (Salario bruto mensual) con el componente 2, por último, A2 (Relación del trabajo) y A3 (Utilidad del programa de educación superior en el trabajo) con el componente 3.

Todas las cargas factoriales anteriores son significativas, ya que son superiores de 0.589 (Cuadro 6).

Tabla 6. Matriz de componentes principales no rotada

	Componente		
	1	2	3
A1	-.342	.437	.629
A2	-.234	.699	-.358
A3	-.125	.517	-.645
A4	-.520	.587	.240
A5	.460	.495	.315
A6	.777	.176	.013
A7	.759	.123	-.007
A8	.737	.306	.034

Tabla 7. Matriz de componentes principales rotada (Rotación varimax)

	Componente		
	1	2	3
A1	-.041	<u>.834</u>	-.071
A2	.015	.259	<u>.777</u>
A3	-.002	-.093	<u>.831</u>
A4	-.200	<u>.720</u>	.338
A5	<u>.660</u>	.343	.034
A6	<u>.775</u>	-.181	-.037
A7	<u>.735</u>	-.220	-.056
A8	<u>.794</u>	-.073	.047

Se observa que en el componente 1, las variables A5 a la A8, se relacionan de manera posita, es decir, se espera que al aumentar el valor de una se incrementen la otros. La misma relación positiva, se presenta entre A1 y A4, en el componente 2 y A2 y A3 en el componente 3. El análisis descriptivo bivariado, también sustenta la relación positiva de las variables (Anexo C).

Para fines de interpretación de la solución factorial de tres componentes principales, el componente 1 se interpretará como las diferencias entre las expectativas del graduado y las características reales de su trabajo actual; el componente 2: El salario indicador principal de la satisfacción laboral; y el componente 3: El valor del conocimiento y las habilidades aprendidas en el programa de educación superior en la actividad laboral del graduado.

El software SPSS 17, calcula las puntuaciones de los factores de cada graduado, los cuales definen, los factores que se tendrán en cuenta en el análisis de clasificación.

La solución factorial de tres componentes principales, tiene un buen ajuste, a pesar de que existen 13 residuos de las 28 diferencias entre las correlaciones observadas y las correlaciones reproducidas superiores a 0.05 (Tabla 8).

Tabla 8. Matriz de correlaciones reproducidas

	A1	A2	A3	A4	A5	A6	A7	A8
A1	.703 ^a	.018	.177	-.0195	-.163	.062	.028	.027
A2	.160	.671 ^a	-.0264	-0,043	-0,021	-.064	-.0124	-0,066
A3	-.137	.621	.699 ^a	0,007	-0,046	-0,072	-0,021	-0,032
A4	.585	.446	.214	.673 ^a	-.141	.028	.075	.047
A5	.257	.125	-.005	.127	.555 ^a	-.147	0,029	-.191
A6	-.180	-.064	-.015	-.298	.448	.635 ^a	-.189	-.023
A7	-.210	-.089	-.027	-.324	.408	.611	.591 ^a	-.191
A8	-.097	.030	.044	-.195	.501	.627	.597	.638 ^a

3.2 FORMACIÓN DE CONGLOMERADOS

El análisis de los coeficientes de aglomeración para la selección de una solución de clasificación, se limita a los coeficientes o el cambio porcentual de los coeficientes de trece a un conglomerado (Tabla 10), teniendo en cuenta que sólo en la etapa 76, se une el conglomerado de miembro único etiquetado con el número 56 (Tabla 9 y Gráfica 1. ver línea roja).

Tabla 9. Esquema de aglomeración por el método Ward

Etapa	Conglomerado que se combina		Coeficientes	conglomerado aparece por		Próxima etapa
	Conglomerado	Conglomerado		Conglomerado	Conglomerado	
	1	2		1	2	
1	45	52	,003	0	0	25
2	3	40	,007	0	0	7
3	48	49	,011	0	0	15
4	25	70	,019	0	0	38
5	36	50	,030	0	0	30
6	16	73	,041	0	0	13
7	3	61	,054	2	0	12
8	28	35	,075	0	0	25
9	1	4	,098	0	0	45
10	9	29	,126	0	0	44
11	18	19	,155	0	0	41
12	3	21	,185	7	0	27
13	16	30	,223	6	0	37
14	53	75	,264	0	0	48

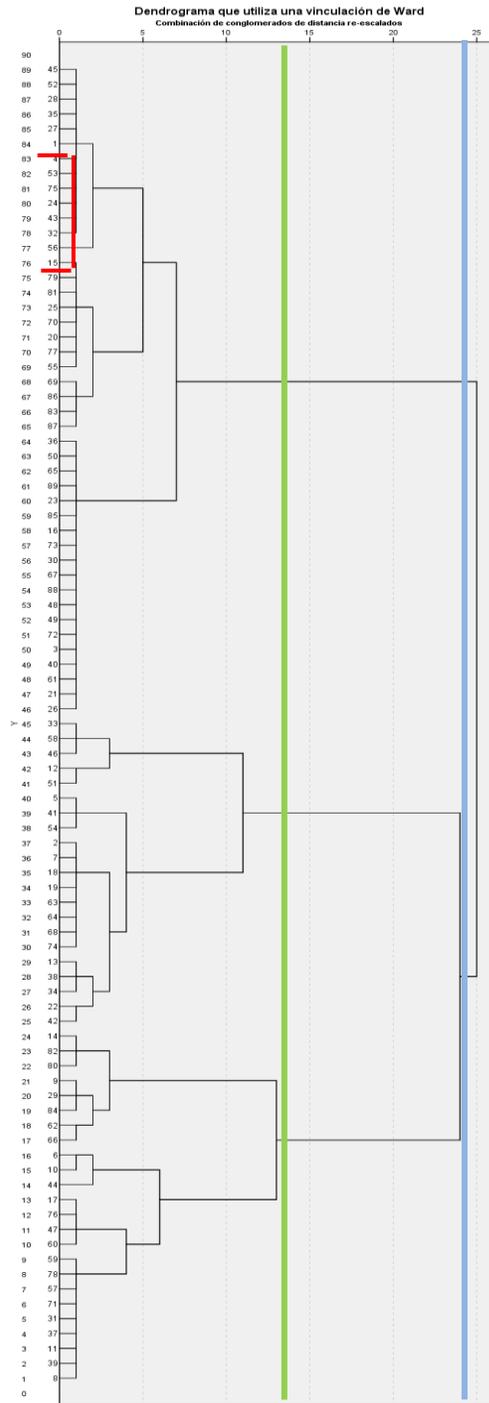
Continuación Tabla 9

Etapa	Conglomerado que se combina		Coeficientes	conglomerado aparece por		Próxima etapa
	Conglomerado 1	Conglomerado 2		Conglomerado 1	Conglomerado 2	
15	48	72	,307	3	0	46
16	57	71	,351	0	0	63
17	23	85	,398	0	0	64
18	15	79	,446	0	0	23
19	67	88	,496	0	0	37
20	20	77	,549	0	0	38
21	69	86	,612	0	0	69
22	11	39	,678	0	0	50
23	15	81	,745	18	0	58
24	24	43	,816	0	0	34
25	28	45	,889	8	1	29
26	65	89	,966	0	0	30
27	3	26	1,045	12	0	46
28	2	7	1,147	0	0	68
29	27	28	1,254	0	25	45
30	36	65	1,366	5	26	71
31	5	41	1,479	0	0	61
32	64	68	1,604	0	0	49
33	14	82	1,736	0	0	59
34	24	32	1,899	24	0	48
35	83	87	2,061	0	0	69
36	31	37	2,227	0	0	55
37	16	67	2,396	13	19	60
38	20	25	2,576	20	4	43
39	17	76	2,778	0	0	51
40	13	38	2,982	0	0	47
41	18	63	3,195	11	0	56
42	59	78	3,428	0	0	70
43	20	55	3,665	38	0	58
44	9	84	3,904	10	0	74
45	1	27	4,158	9	29	65
46	3	48	4,432	27	15	60
47	13	34	4,716	40	0	73
48	24	53	5,048	34	14	65
49	64	74	5,392	32	0	56
50	8	11	5,767	0	22	55

Continuación Tabla 9

Etapa	Conglomerado que se combina		Coeficientes	conglomerado aparece por		Próxima etapa
	Conglomerado 1	Conglomerado 2		Conglomerado 1	Conglomerado 2	
51	17	47	6,161	39	0	66
52	22	42	6,584	0	0	73
53	6	10	7,123	0	0	75
54	33	58	7,693	0	0	62
55	8	31	8,328	50	36	63
56	18	64	8,998	41	49	68
57	62	66	9,690	0	0	74
58	15	20	10,383	23	43	72
59	14	80	11,082	33	0	78
60	3	16	11,877	46	37	64
61	5	54	12,778	31	0	81
62	33	46	13,822	54	0	79
63	8	57	14,889	55	16	70
64	3	23	15,988	60	17	71
65	1	24	17,151	45	48	76
66	17	60	18,349	51	0	80
67	12	51	19,643	0	0	79
68	2	18	21,078	28	56	77
69	69	83	22,658	21	35	72
70	8	59	24,640	63	42	80
71	3	36	26,634	64	30	84
72	15	69	28,969	58	69	82
73	13	22	31,328	47	52	77
74	9	62	34,567	44	57	78
75	6	44	38,026	53	0	83
76	1	56	41,818	65	0	82
77	2	13	46,842	68	73	81
78	9	14	52,310	74	59	86
79	12	33	58,765	67	62	85
80	8	17	65,347	70	66	83
81	2	5	72,825	77	61	85
82	1	15	81,687	76	72	84
83	6	8	93,857	75	80	86
84	1	3	108,981	82	71	88
85	2	12	131,552	81	79	87
86	6	9	157,807	83	78	87
87	2	6	209,567	85	86	88
88	1	2	264,000	84	87	0

Gráfica 1. Dendrograma por el método Ward



El mayor cambio porcentual en el coeficiente de aglomeración se presenta al pasar de tres conglomerados a dos (Tabla 10). Podrían ser estas, dos soluciones posibles de clasificación. Si embargo, se prevé que la solución de tres conglomerados obtenga mejores resultados, ya que examinado la distancia de corte del último conglomerado para la formación de dos conglomerados, es más tardía (distancia 24, línea azul), que en el caso de tres conglomerados (distancia 13, líneas verde) (Gráfica 1).

Tabla 10. Coeficientes de aglomeración

Número de conglomerados	Coeficiente de aglomeración	Cambio porcentual en el coeficiente del nivel siguiente
13	41,818	12,0%
12	46,842	11,7%
11	52,310	12,3%
10	58,765	11,2%
9	65,347	11,4%
8	72,825	12,2%
7	81,687	14,9%
6	93,857	16,1%
5	108,981	20,7%
4	131,552	20,0%
3	157,807	32,8%
2	209,567	26,0%
1	264,000	

En la solución de clasificación de dos conglomerados, hay diferencias significativas entre la media de la diferencia entre las expectativas del graduado y las características reales de su trabajo actual (componente principal 1) (Tabla 11). Además, en el valor del conocimiento y las habilidades aprendidas en el programa de educación superior en la actividad laboral del graduado (componente principal 3) (Tabla 13). Pero no se detectan diferencias significativas en los dos conglomerados, respecto a la media del salario indicador principal de la satisfacción laboral (componente principal 2) (Tabla 12).

En cambio, en la solución de tres conglomerados, si se detentan diferencias en las medias de las tres variables anteriores, a un nivel de significancia del 5% (Tabla 14). Por lo tanto se, concluye que es mejor agrupar a los graduados en tres conglomerados.

Tabla 11. Prueba T para la igualdad de medias componente principal 1 entre dos conglomerados

Prueba T para la igualdad de medias							
						para la diferencia	
	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	Inferior	Superior
Componente principal 1	-4,420	87	,000	-,85170127	,19268700	-1,23468755	-,46871498

Tabla 12. Prueba T para la igualdad de medias componente principal 2 entre dos conglomerados

Prueba T para la igualdad de medias							
						para la diferencia	
	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	Inferior	Superior
Componente principal 2	,544	87	,588	,11569868	,21286684	-,30739722	,53879457

Tabla 13. Prueba U de Mann-Whitney para la igualdad de medias componente principal 3 entre dos conglomerados*

Estadísticos de contraste	Componente principal 3
U de Mann-Whitney	186,000
Sig. asintót. (bilateral)	,000

* se aplica la prueba no paramétrica por falta de cumplimiento del supuesto de distribución normal de las observaciones.

Tabla 14. Anova: Solución tres conglomerados

		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Componente principal 1	Inter-grupos	30.850	2	15.425	23.212	.000
	Intra-grupos	57.150	86	.665		
	Total	88.000	88			
Componente principal 2	Inter-grupos	33.339	2	16.670	26.227	.000
	Intra-grupos	54.661	86	.636		
	Total	88.000	88			
Componente principal 3	Inter-grupos	42.004	2	21.002	39.268	.000
	Intra-grupos	45.996	86	.535		
	Total	88.000	88			

3.3 VALIDACIÓN DE LA SOLUCIÓN DE TRES CONGLOMERADOS

La validez de la solución de clasificación se evaluó aplicando el método no jerárquico umbral paralelo donde los centros iniciales son los obtenidos por el método jerárquico Ward.

Los resultados comparativos de la pertenencia a los conglomerados por los dos métodos (Tabla 15), confirman la consistencia y la estabilidad de los resultados, es decir, la aceptación práctica y teórica de la solución obtenida por el método jerárquico Ward, teniendo en cuenta que hay consistencia en los graduados agrupados por los dos métodos. Salvo por seis casos, pero son comparables los tamaños y los casos de los conglomerados por ambos métodos.

Además las varianzas entre los conglomerados por los dos métodos son comparables, si se contrastan los cuadrados medios del error (Tabla 16 y Tabla 17).

Al igual que en la solución por el método jerárquico hay diferencias significativas entre todas las variables, con un nivel de significancia del 5% (Tabla 16).

Tabla15. Pertenencia a los conglomerados

NÚMERO DE CASO	MÉTODO JERÁRQUICO WARD (CONGLOMERADO)	MÉTODO NO JERARQUICO UMBRAL PARALELO (CONGLOMERADO)	NÚMERO DE CASO	MÉTODO JERÁRQUICO WARD (CONGLOMERADO)	MÉTODO NO JERARQUICO UMBRAL PARALELO (CONGLOMERADO)
1	1	1	50	2	2
2	1	1	51	2	2
3	1	1	52	2	2
4	3	1	53	2	2
5	3	1	54	2	2
6	3	1	55	2	2
7	1	1	56	2	2
8	1	1	57	2	2
9	1	1	58	2	2
10	1	1	59	2	2
11	1	1	60	2	2
12	1	1	61	2	2
13	1	1	62	2	2
14	1	1	63	2	2
15	1	1	64	2	2
16	1	1	65	2	2
17	3	1	66	2	2
18	1	1	67	2	2
19	1	1	68	2	2
20	1	1	69	2	2
21	1	1	70	3	3
22	1	1	71	3	3
23	1	1	72	3	3
24	1	1	73	3	3
25	1	1	74	3	3
26	1	1	75	3	3
27	1	1	76	3	3
28	1	1	77	3	3
29	1	1	78	3	3
30	1	1	79	3	3
31	1	1	80	3	3
32	1	1	81	3	3
33	1	1	82	3	3
34	1	1	83	2	3
35	1	1	84	3	3
36	1	1	85	3	3
37	1	1	86	3	3
38	1	1	87	3	3
39	1	1	88	3	3
40	1	1	89	3	3
41	1	1			
42	1	1			
43	1	1			
44	3	1			
45	1	1			
46	1	1			
47	1	1			
48	1	1			
49	1	1			

Tabla 16. Anova Método Umbral Paralelo

		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Componente principal 1	Inter-grupos	29.345	2	14.672	21.513	.000
	Intra-grupos	58.655	86	.682		
	Total	88.000	88			
Componente principal 2	Inter-grupos	31.081	2	15.541	23.481	.000
	Intra-grupos	56.919	86	.662		
	Total	88.000	88			
Componente principal 3	Inter-grupos	50.805	2	25.403	58.734	.000
	Intra-grupos	37.195	86	.432		
	Total	88.000	88			

Tabla 17. Anova Método Ward

		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Componente principal 1	Inter-grupos	30.850	2	15.425	23.212	.000
	Intra-grupos	57.150	86	.665		
	Total	88.000	88			
Componente principal 2	Inter-grupos	33.339	2	16.670	26.227	.000
	Intra-grupos	54.661	86	.636		
	Total	88.000	88			
Componente principal 3	Inter-grupos	42.004	2	21.002	39.268	.000
	Intra-grupos	45.996	86	.535		
	Total	88.000	88			

3.4 INTERPRETACIÓN DE LA SOLUCIÓN DE TRES CONGLOMERADOS

Se observa en el gráfico de medias (Gráfica 2) que no hay diferencias entre las medias del conglomerado 1 y 3, respecto a las diferencias entre las expectativas del graduado y las características reales de su trabajo actual (Componente 1), con un nivel de significancia del 0,05 (Tabla 18. Los asteriscos indican diferencias).

En el grupo de graduados 1 y 3, el valor negativo en la gráfica 2, indica que en estos dos grupos hay menos diferencias entre las expectativas del graduado y las características encontradas en su trabajo actual. En cambio sucede lo contrario en el grupo de graduados 2.

Gráfica 2. Media de los valores del componente de regresión 1

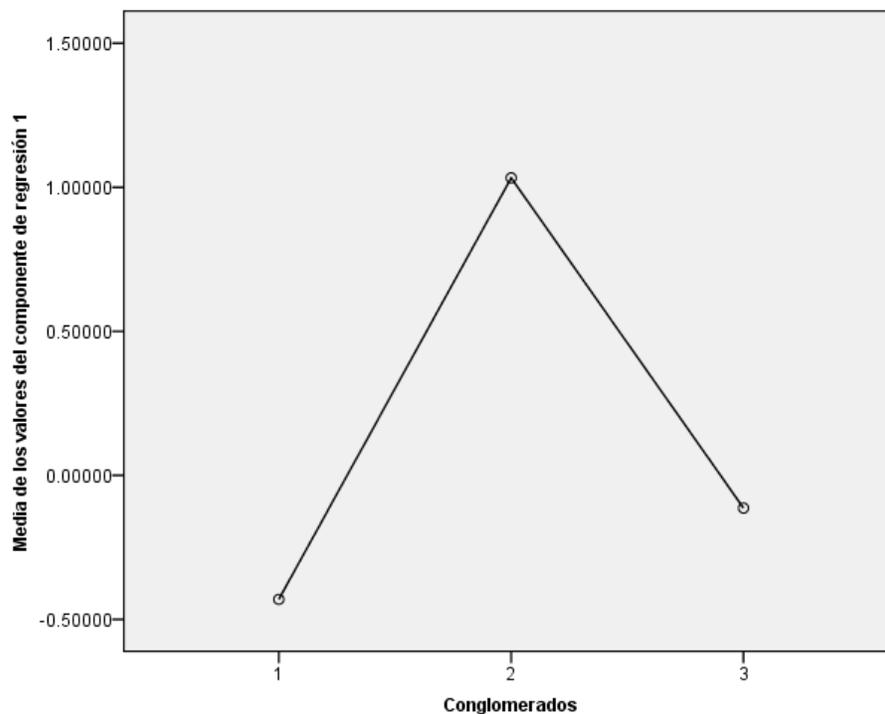


Tabla 18. Comparaciones múltiples – Diferencia Mínima Significativa (DMS)

Variable dependiente: Valores del componente de regresión 1

	(I) Ward Method	(J) Ward Method	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
						Límite inferior	Límite superior
DMS	1	2	-1.46296447 [*]	,21621155	,000	-1,8927788	-1,0331501
		3	-,31684597	,20686199	,129	-,7280740	,0943821
	2	1	1.46296447 [*]	,21621155	,000	1,0331501	1,8927788
		3	1.14611850 [*]	,24358424	,000	,6618891	1,6303479
	3	1	,31684597	,20686199	,129	-,0943821	,7280740
		2	-1.14611850 [*]	,24358424	,000	-1,6303479	-,6618891

Con relación al salario bruto indicador principal de la satisfacción laboral (Componente 2), hay diferencias significativas en las medias de los tres grupos (Gráfica 3 y Tabla 19), con un nivel de significancia del 5%.

Se observa en la gráfica de medias, que el grupo 3 de graduados tiene los salarios más bajos, el grupo 2 los más altos y el grupo 1 devenga salarios intermedios entre los grupos 2 y 3.

Gráfica 3. Media de los valores del componente de regresión 2

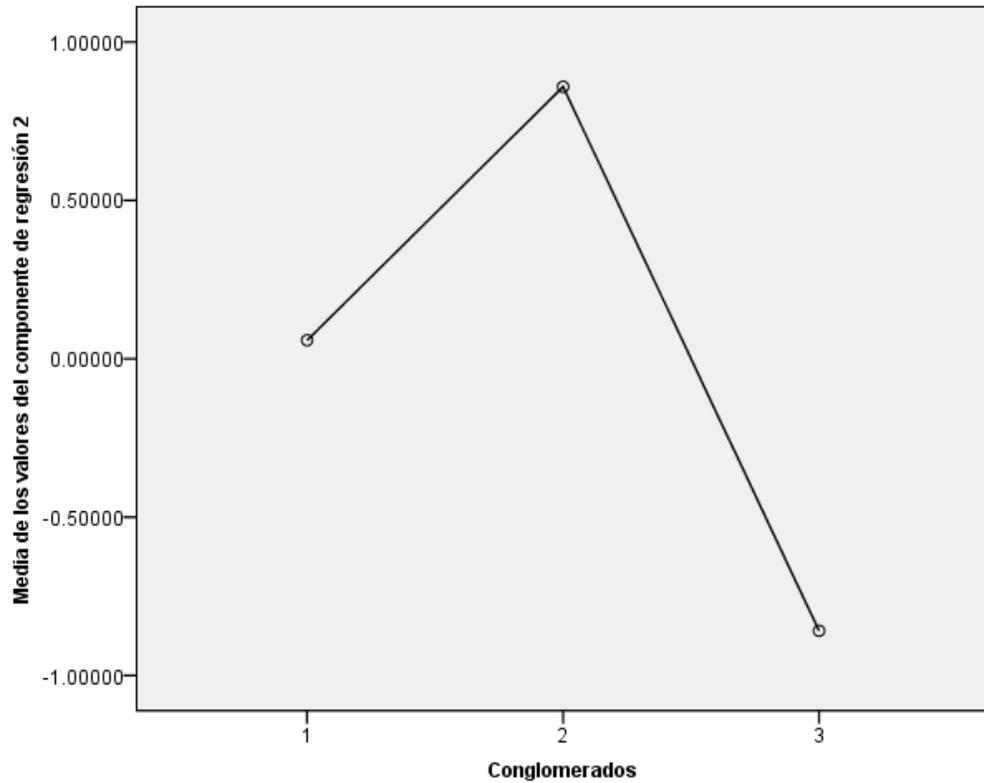


Tabla 19. Comparaciones múltiples – Diferencia Mínima Significativa (DMS)

Variable dependiente: Valores del componente de regresión 2

	(I) Ward Method	(J) Ward Method	Diferencia de medias (I-J)	Error típico	Sig.	95%	
						Límite inferior	Límite superior
DMS	1	2	-.80034798	,21145125	,000	-1,2206991	-,3799968
		3	,91723950	,20230753	,000	,5150654	1,3194136
	2	1	,80034798	,21145125	,000	,3799968	1,2206991
		3	1,71758747	,23822128	,000	1,2440192	2,1911557
	3	1	-.91723950	,20230753	,000	-1,3194136	-,5150654
		2	-1,71758747	,23822128	,000	-2,1911557	-1,2440192

El valor del conocimiento y las habilidades aprendidas en el programa de educación superior en el trabajo, es significativamente diferente en cada uno de los grupo de graduados, a un nivel del 5% (Gráfica 4 y Tabla 20).

Donde el valor que imprime el conocimiento y las habilidades aprendidas del graduado en el programa de estudio que culminó es mayor en el grupo 1, y va descendiendo la trascendencia del valor de los conocimientos y las habilidades en la aplicación en el trabajo (Gráfica 4).

Gráfica 4. Media de los valores del componente de regresión 3

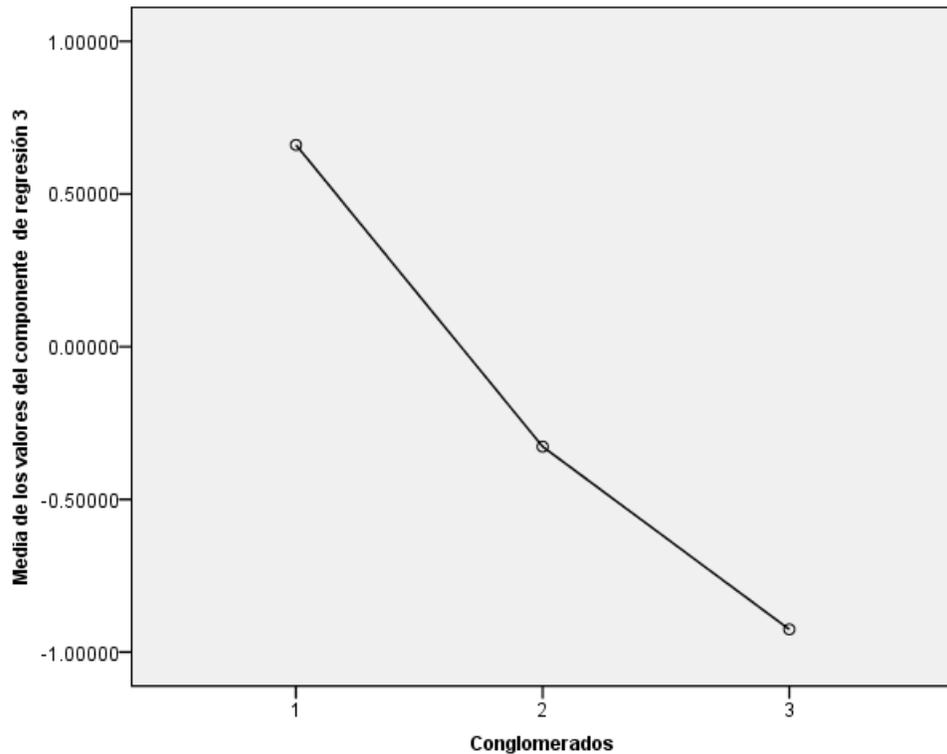


Tabla 20. Comparaciones múltiples – Diferencia Mínima Significativa (DMS)

Variable dependiente: Valores del componente de regresión 3

	(I) Ward Method	(J) Ward Method	Diferencia de medias (I-J)	Error típico	Sig.	95%	
						Límite inferior	Límite superior
DMS	1	2	.98787455 [*]	,19396884	,000	,6022773	1,3734718
		3	1.58601336 [*]	,18558111	,000	1,2170903	1,9549364
	2	1	-.98787455 [*]	,19396884	,000	-1,3734718	-,6022773
		3	.59813881 [*]	,21852557	,008	,1637243	1,0325533
	3	1	-1.58601336 [*]	,18558111	,000	-1,9549364	-1,2170903
		2	-.59813881 [*]	,21852557	,008	-1,0325533	-,1637243

Lo anterior resume que las variables, identifican diferentes características del empleo de los graduados en los tres conglomerados. Por lo tanto, se observa en la gráfica de individuos en conglomerados, agrupaciones definidas y disímiles entre sí (Gráfica 5). También confirma la formación de agrupaciones definidas, las medias de cada conglomerado (Tabla 21), donde el centroide más bajo según las coordenadas (medias de cada variable) es el del conglomerado 3, luego en un punto intermedio el conglomerado 2, y el centroide más alto respecto al eje componente 3 es el conglomerado 1.

Gráfica 5. Gráfica de individuos en conglomerados

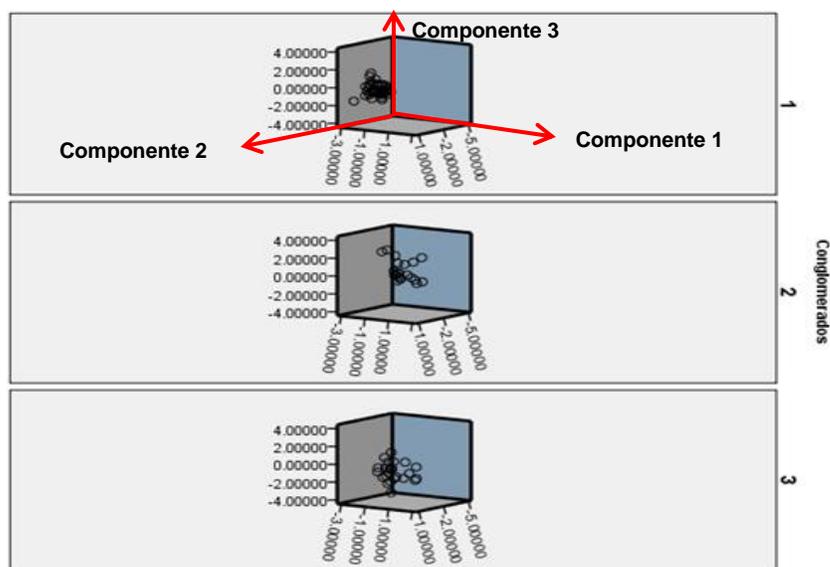


Tabla 21. Estadísticos descriptivos de los tres conglomerados

Conglomerados	N	Media	Desviación típica	la media al 95%		Mínimo	Máximo	
				Límite inferior	Límite superior			
Componente principal 1	1	44	-,4306355	,69088981	-,6406852	-,2205858	-,241435	,75421
	2	21	1,0323290	,93949605	,6046755	1,4599825	-,108183	2,81981
	3	24	-,1137895	,90821431	-,4972948	,2697158	-,141168	1,71486
	Total	89	,0000000	1,00000000	-,2106523	,2106523	-,241435	2,81981
Componente principal 2	1	44	,0584993	,62023035	-,1300679	,2470666	-,151099	1,74750
	2	21	,8588473	1,03984502	,3855155	1,3321791	-,38280	3,12454
	3	24	-,8587402	,84683457	-,12163270	-,5011533	-,300074	,93378
	Total	89	,0000000	1,00000000	-,2106523	,2106523	-,300074	3,12454
Componente principal 3	1	44	,6607830	,27449655	,5773284	,7442376	-,48724	,99924
	2	21	-,3270916	,78612186	-,6849299	,0307468	-,244710	,80913
	3	24	-,9252304	1,14960049	-,14106641	-,4397967	-,405050	,76992
	Total	89	,0000000	1,00000000	-,2106523	,2106523	-,405050	,99924

3.5 TIPOLOGÍA DE LA SOLUCIÓN DE TRES CONGLOMERADOS

Una vez observadas las tendencias en los conglomerados respecto a las componentes principales o las variables nuevas (numeral 3.5): Diferencias en las expectativas del graduado y las características reales de su trabajo actual, el salario indicador principal de la satisfacción laboral y el valor del conocimiento y las habilidades aprendidas en el programa de educación superior en la actividad laboral del graduado. Se estudian las variables iniciales respecto a los conglomerados, para definir características en cada uno de ellos.

Conglomerado 1: “Empleo trascendente”

Está conformado por graduados que perciben un salario promedio de un millón quinientos treinta mil (\$1.530.000) con una desviación de quinientos veintidós mil pesos (\$522.702) (Tabla 22). Consideran que su trabajo está altamente relacionado con el programa de educación superior, por lo tanto, han sido útiles los conocimientos y habilidades aprendidas en la universidad (Tabla 23 y 24).

Las diferencias entre las expectativas del graduado y la importancia que le da la empresa a los factores: estabilidad laboral, mejora de los ingresos laborales, la

oportunidad de adquirir nuevos conocimientos y tener la posibilidad de afrontar nuevos retos, están en concordancia, es decir, predomina la no diferencia entre la importancia graduado-empresa o las diferencias son leves (Tabla 25, 27, 26 y 28). Además esa diferencia es positiva, en el sentido de que si es importante para el graduado es también importante para la empresa.

En general, los graduados se encuentran satisfechos con el trabajo actual (Tabla 29).

Conglomerado 2: “Empleo con aprendizaje en la empresa”

Está conformado por egresados con el mayor promedio salarial, dos millones cien mil pesos (\$2.100.238) y una desviación de quinientos sesenta mil (\$561.959) (Tabla 22).

Se destaca en este grupo graduados laborando en empresas en las cuales los cargos que ocupan se encuentran medianamente relacionados con el título profesional recibido y han sido medianamente útiles en el trabajo los conocimientos, habilidades y destrezas aprendido en el programa académico que terminaron (Tabla 23 y 24).

Es el grupo de graduados que manifiesta menos afinidad entre la importancia dada a la estabilidad laboral y a la mejora de los ingresos y la considerada por la empresa. Aunque para la empresa en general, estos aspectos son medianamente importantes, para el graduado son más cruciales (Tabla 25 y 27).

La mayor proporción de graduados pueden adquirir nuevos conocimientos y afrontar nuevos retos profesionales en sus empleos (Tabla 28 y 28).

En general, están satisfechos con el empleo actual (Tabla 29).

Conglomerado 3: “Empleo sin valor”

Está integrado por los egresados que devengan los ingresos promedios más bajos (\$1.070.356), con una desviación de cuatrocientos veinte mil pesos (\$422.439) (Tabla 22).

Son el grupo de graduados con los cargos menos relacionados con el programa académico culminado y de la misma manera han sido menos útiles los conocimientos, habilidades y destrezas aprendidas durante la etapa de formación en la universidad (Tabla 23 y 24).

Las diferencias entre las expectativas del graduado y la importancia que le da la empresa a los factores: estabilidad laboral, la oportunidad de adquirir nuevos conocimientos y tener la posibilidad de afrontar nuevos retos, están en concordancia, es decir, predomina la no diferencia entre la importancia graduado-empresa o las diferencias son leves (Tabla 25, 26 y 28). Además esa diferencia es positiva, en el sentido de que si es importante para el graduado es también importante para la empresa, excepto en la mejora de los ingresos laborales (Tabla 27).

Son el grupo de graduados menos satisfechos con su situación laboral, es decir, medianamente satisfechos (Tabla 29).

Tabla 22. Estadísticos descriptivos Salario mensual

Método Ward	N	Media	Desv. típ.
1 A1 (Salario mensual)	44	1530772,73	522702,319
2 A1 (Salario mensual)	21	2100238,10	561959,558
3 A1 (Salario mensual)	24	1070356,00	422439,676

Tabla 23. Tabla de contingencia variable A2 vs conglomerados

% dentro de Método Ward

A2. Relación del trabajo con el programa de educación que término	Método Ward			Total
	1	2	3	
1			8,3%	2,2%
2		4,8%	8,3%	3,4%
3		19,0%	50,0%	18,0%
4	100,0%	76,2%	33,3%	76,4%
Total	100,0%	100,0%	100,0%	100,0%

Tabla 24. Tabla de contingencia variable A3 vs conglomerados

% dentro de Método Ward

A3. Utilidad del programa de educación superior en el trabajo	Método Ward			Total
	1	2	3	
1			4,2%	1,1%
2			8,3%	2,2%
3	4,5%	57,1%	50,0%	29,2%
4	95,5%	42,9%	37,5%	67,4%
Total	100,0%	100,0%	100,0%	100,0%

Tabla 25. Tabla de contingencia variable A5 vs conglomerados

% dentro de Método Ward

A5. Diferencia entre la importancia dada por el graduado y la empresa a la estabilidad laboral	Método Ward			Total
	1	2	3	
-2			4,2%	1,1%
-1	4,5%			2,2%
0	54,5%	4,8%	58,3%	43,8%
1	27,3%	28,6%	33,3%	29,2%
2	13,6%	61,9%	4,2%	22,5%
3		4,8%		1,1%
Total	100,0%	100,0%	100,0%	100,0%

Tabla 26. Tabla de contingencia variable A6 vs conglomerados

% dentro de Método Ward

A6. Diferencia entre la importancia dada por el graduado y la empresa a la oportunidad de adquirir nuevos conocimientos	Método Ward			Total
	1	2	3	
-1	2,3%			1,1%
0	54,5%	19,0%	33,3%	40,4%
1	34,1%	28,6%	50,0%	37,1%
2	9,1%	38,1%	16,7%	18,0%
3		14,3%		3,4%
Total	100,0%	100,0%	100,0%	100,0%

Tabla 27. Tabla de contingencia variable A7 vs conglomerados

% dentro de Método Ward

A7. Diferencia entre la importancia dada por el graduado y la empresa a la mejora de los ingresos laborales.	Método Ward			Total
	1	2	3	
-1	2,3%			1,1%
0	34,1%	4,8%	25,0%	24,7%
1	36,4%	19,0%	20,8%	28,1%
2	25,0%	52,4%	33,3%	33,7%
3	2,3%	9,5%	12,5%	6,7%
4		14,3%	8,3%	5,6%
Total	100,0%	100,0%	100,0%	100,0%

Tabla 28. Tabla de contingencia variable A8 vs conglomerados

% dentro de Método Ward

A8. Diferencia entre la importancia dada por el graduado y la empresa a la posibilidad de afrontar nuevos retos	Método Ward			Total
	1	2	3	
-2	2,3%			1,1%
0	52,3%	28,6%	50,0%	46,1%
1	29,5%	28,6%	20,8%	27,0%
2	15,9%	28,6%	25,0%	21,3%
3		9,5%	4,2%	3,4%
4		4,8%		1,1%
Total	100,0%	100,0%	100,0%	100,0%

Tabla 29. Tabla de contingencia variable A8 vs conglomerados

% dentro de Método Ward

A4. Satisfacción con el trabajo	Método Ward			Total
	1	2	3	
1			4,2%	1,1%
2			8,3%	2,2%
3	2,3%	19,0%	54,2%	20,2%
4	63,6%	47,6%	25,0%	49,4%
5	34,1%	33,3%	8,3%	27,0%
Total	100,0%	100,0%	100,0%	100,0%

3.6 OBTENCIÓN DE PERFILES DE LOS TRES CONGLOMERADOS SOBRE VARIABLES ADICIONALES

No hay asociación o relación entre la solución de clasificación de tres conglomerados y las variables género, nivel del cargo (desde el punto de vista del nivel de responsabilidad en la definición de los objetivos de la empresa) y tipo de contrato (Tabla 30, 31 y 32), a un nivel de significancia del 5% ($p = 0.1775 > \alpha = 0.025$; $p = 0.2445 > \alpha = 0.025$; $p = 0.455 > \alpha = 0.025$, respectivamente).

Tabla 30. Tabla de contingencia variable género vs conglomerados

% dentro de Método Ward

Género	Método Ward			Total
	1	2	3	
Masculino	68,2%	52,4%	54,2%	60,7%
Femenino	31,8%	47,6%	45,8%	39,3%
Total	100,0%	100,0%	100,0%	100,0%

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	2.071	2	,355
N de casos válidos	89		

Tabla 31. Tabla de contingencia variable nivel de responsabilidad en la definición de los objetivos de la empresa vs conglomerados

% dentro de Método Ward

Nivel de responsabilidad en la definición de los objetivos de empresa	Método Ward			Total
	1	2	3	
Nada de responsabilidad	20,5%	28,6%	37,5%	27,0%
Poca responsabilidad	25,0%	33,3%	20,8%	25,8%
Alta responsabilidad	54,5%	38,1%	41,7%	47,2%
Total	100,0%	100,0%	100,0%	100,0%

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	3.428	4	,489
N de casos válidos	89		

Tabla 32. Tabla de contingencia variable contrato vs conglomerados

% dentro de Método Ward

Contrato	Método Ward			Total
	1	2	3	
A término definido	66,7%	66,7%	72,2%	68,1%
A término indefinido	33,3%	33,3%	27,8%	31,9%
Total	100,0%	100,0%	100,0%	100,0%

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	,189	2	,910
N de casos válidos	69		

Con relación a las variables escuela y facultad, cada una contrastada con la solución de clasificación, no se observa asociación con un conglomerado en particular¹³ (Tabla 33 y 34). En cambio, en referencia al tamaño de la empresa, los graduados del conglomerado dos que perciben un promedio salarial mayor que los otros grupos, trabajan en mayor proporción en grandes empresas (Tabla 35).

¹³ No se aplica la prueba de Chi-cuadrado, ya que más del 20% de las celdas tienen una frecuencia esperada inferior a 5.

Tabla 33. Tabla de contingencia variable escuela vs conglomerados

% dentro de Método Ward

Escuela	Método Ward			Total
	1	2	3	
Humanidades	54,5%	38,1%	29,2%	43,8%
Ingenierías	43,2%	61,9%	62,5%	52,8%
Administración	2,3%		8,3%	3,4%
Total	100,0%	100,0%	100,0%	100,0%

Tabla 34. Tabla de contingencia variable facultad vs conglomerados

% dentro de Método Ward

Facultad	Método Ward			Total
	1	2	3	
Ing. Industrial	6,8%	4,8%	12,5%	7,9%
Ing. Mecánica	6,8%	4,8%	12,5%	7,9%
Ing. Electrónica		23,8%	12,5%	9,0%
Psicología	36,4%	28,6%	16,7%	29,2%
Ing. Ambiental	13,6%	14,3%	20,8%	15,7%
Comunicación social y Periodismo	18,2%	9,5%	12,5%	14,6%
Administración de empresas	2,3%		8,3%	3,4%
Ing. Civil	13,6%	14,3%	4,2%	11,2%
Ing. Informática	2,3%			1,1%
Total	100,0%	100,0%	100,0%	100,0%

Tabla 35. Tabla de contingencia variable tamaño empresa vs conglomerados

% dentro de Método Ward

Tamaño empresa	Método Ward			Total
	1	2	3	
Micro	6,8%	9,5%	16,7%	10,1%
Pequeña	40,9%	14,3%	45,8%	36,0%
Mediana	15,9%	19,0%	12,5%	15,7%
Grande	36,4%	57,1%	25,0%	38,2%
Total	100,0%	100,0%	100,0%	100,0%

CONCLUSIONES

Los resultados derivados del seguimiento a graduados, mediante un análisis estadístico multivariado de clasificación, permite analizar la situación laboral de los graduados de uno a dos años después de recibir el título profesional y agruparlos de la siguiente forma:

- El 49% de los graduados de la UPB de uno a dos años después del grado, tienen **un empleo trascendente** (Conglomerado 1), debido a que los factores considerados por los egresados importantes con relación a las características de un buen empleo; estabilidad laboral, mejora de los ingresos laborales, la oportunidad de adquirir nuevos conocimientos y tener la posibilidad de afrontar nuevos retos, se encuentra en concordancia con la importancia dada por empresa. Es decir, no causa malestar en el graduado su situación laboral por que los factores que definen un buen empleo para ellos están en sintonía con la empresa donde se desempeñan

Adicionalmente, sus cargos están muy relacionados con el programa académico que culminaron y ha sido muy útil lo aprendido en el desarrollo en las actividades propias del cargo. En general están satisfechos con el trabajo actual.

- En el **empleo con aprendizaje en la empresa** (24% de los graduados) (Conglomerado 2), es el grupo donde la mayor proporción de graduados pueden adquirir nuevos conocimientos y afrontar nuevos retos profesionales en sus empleos.

Se destaca en este grupo graduados laborando en empresas en las cuales los cargos que ocupan se encuentran medianamente relacionados con el título profesional recibido y han sido medianamente útiles en el trabajo los

conocimientos, habilidades y destrezas aprendido en el programa académico que terminaron.

Aunque es el grupo con el mayor promedio salarial, dos millones cien mil pesos (\$2.100.238) y una desviación de quinientos sesenta mil (\$561.959), manifiestan, que es crucial la mejora de sus ingresos laborales, al igual que la estabilidad laboral en consideración con la importancia que le da la empresa a estos aspectos, que es medianamente importante. Las empresas donde trabajan la mayor proporción de estos graduados son en grandes empresas.

En general, están satisfechos con el empleo actual.

Es importante resaltar que estos tres perfiles del empleo de los graduados, demuestran ser independientes del género, la escuela, la facultad a la cual perteneció, el nivel del cargo desde el punto de vista de la responsabilidad en la definición de los objetivos de la empresa y el tipo de contrato, es decir, que al interior de cada conglomerado se encuentra diversidad de cada uno de estos aspectos y similarmente proporcionales de un conglomerado a otro, por lo tanto en esta etapa de transición a la vida laboral, un año a dos años después del grado, priman otras características diferenciadoras del empleo o de la vida laboral como son el salario, el trabajo relacionado con la carrera, la utilidad de los conocimientos y habilidades aprendidas en la universidad en el trabajo, la satisfacción laboral, la búsqueda de mejoras en los ingresos, la estabilidad laboral, la posibilidad de adquirir nuevos conocimientos y afrontar nuevos retos.

RECOMENDACIONES

La estructura conseguida con el análisis de clasificación con las variables ordinales tratadas como variables cualitativas por su escalamiento en forma de escala de Likert, permite analizar el perfil de cada uno de los conglomerados respecto a las características que los definen de manera consistente. Pero se desea seguir enriqueciendo el trabajo de clasificación comparando el análisis de clasificación logrado en el presente estudio, con una solución de clasificación empleando el coeficiente de Gower como medida de similitud. El cual permite el uso de variables mixtas (tanto cualitativas como cuantitativas).

Pero ello requiere un mayor tiempo para el análisis, ya el coeficiente de Gower es una subrutina del programa R. La subrutina es:

```
daisy (x, metric = c("euclidean", "manhattan", "gower"),
```

```
stand = FALSE, type = list(), weights = rep.int(1, p))
```

BILIOGRAFÍA

AGENCIA NACIONAL DE EVALUACIÓN DE LA CALIDAD Y ACREDITACIÓN- ANECA, El profesional flexible en la sociedad del conocimiento (Proyecto REFLEX), Madrid, 2007. 72 p.

CHAVÉZ, Dunia, Utilización del Análisis de Cluster con Variables Mixtas en la Selección de Genotipos de maíz (zea mays), Monografía, 2010, Cuba: Universidad Agraria de la Habana.

DÍAZ MONROY Luis Guillermo, Estadística multivariada: Inferencia y métodos, Bogotá: Universidad Nacional de Colombia, 2000, Capítulo 4.

DÍAZ MONROY Luis Guillermo; MORALES RIVERA Mario Alfonso, Análisis estadístico de datos categóricos, 2009, Bogotá: Universidad nacional de Colombia, Primera edición, Capítulo 1 y 2.

EVERITT, Brian, Cluster Analysis, 1974, First published, Published on behalf on the social Social Science Research Council by Hernemann, Capítulo 2 y 3.

GUERRA, Prieto, Técnicas Estadísticas de Clasificación, Monografía, 2006, México: Universidad Autónoma del Estado de Hidalgo.

HAIR Joseph F; ANDERSON Rolph E, Análisis Multivariante, Quinta edición, Madrid: Prentice Hall, 2004, Capítulo 9.

MAEHLER, Martin, Package 'Cluster' Software R, 2012.

MALHOTRA Naresh K, Investigación de mercados, Quinta edición, México: Prentice Hall, 2008, Capítulo 15 y capítulo 20.

MINISTERIO DE EDUCACIÓN NACIONAL DE COLOMBIA, Estudio de seguimiento a graduados, Bogotá: Ministerio de Educación Nacional, 2008. 113 p.

MORA José Ginés; CAROT José Miguel; CONCHADO Andrea, El profesional flexible en la sociedad del conocimiento (Proyecto PROFLEX), Madrid, 2008. 74 p.

SPSS inc, Manual de SPSS 17.0, 2007, Chicago.

TEICHLER Ulrich, Higher Education and Graduate Employment in Europe (Cheers Project), Alemania: Gesamthochschule Kassel University, 1997. 85 p.

ANEXOS

ANEXO A. Análisis descriptivo univariado

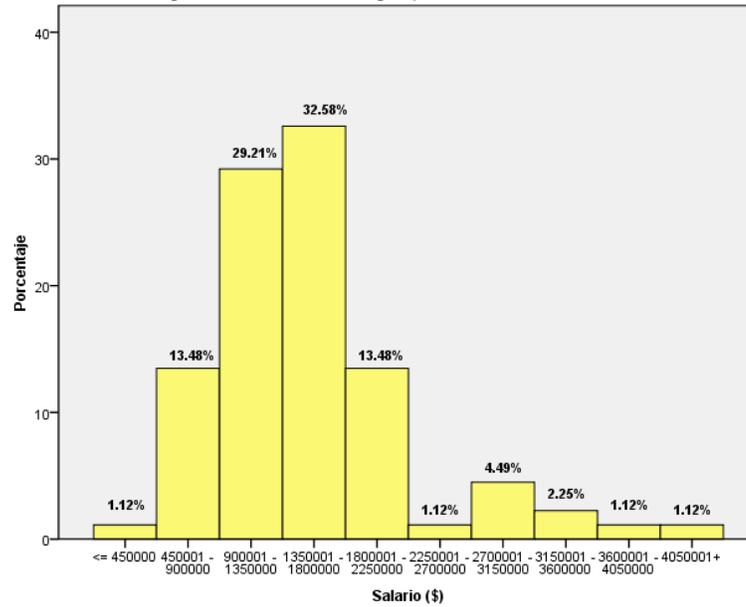
A1. Salario bruto mensual

El promedio salarial de los graduados es de \$1.540.983 pesos con una desviación de \$724.949 pesos (Tabla1). Esta desviación considerable, permite prever la existencia de grupos de graduados notoriamente diferentes por niveles salariales. Lo cual se observa en el histograma (Gráfica 1), donde los salarios de los graduados se encuentran sesgados a la derecha, es decir, existen salarios no convencionales dentro de los graduados (\$2.700.000 en adelante). Además hay dos grupos de salarios que son predominantes entre los graduados y se encuentran entre \$900.000 y \$1.800.00 pesos (61.79%). Por último se destaca un grupo de graduados por debajo de los \$900.000 pesos (14.6%).

Tabla 1. Estadísticos descriptivos

N	Válidos	89
	Perdidos	0
Media		1540983.64
Mediana		1426000.00
Moda		2000000
Desv. típ.		724949.763
Mínimo		450000
Máximo		4500000
Percentiles	25	1000000.00
	50	1426000.00
	75	1800000.00

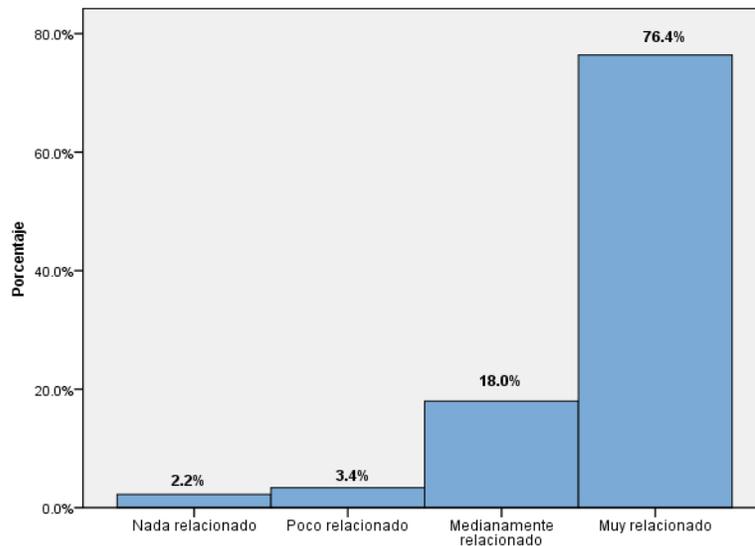
Gráfica 1. Histograma de datos agrupados salario bruto mensual



A2. Relación del trabajo con el programa de educación superior culminado.

La mayor proporción de graduados consideran que el trabajo su trabajo actual está muy relacionado con el programa que culminó (76,4%), Se observa un grupo de graduados que reconocen que su trabajo actual no guarda relación con los estudios superiores realizados (5,6%) (Gráfica 2).

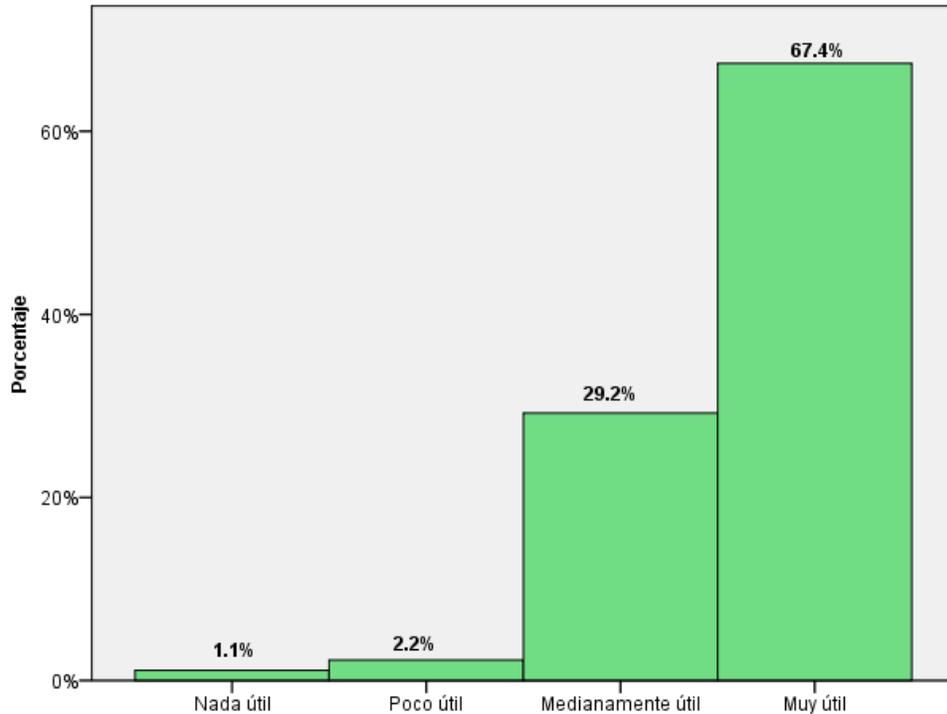
Gráfica 2. Relación del trabajo con el programa culminado



A3. Utilidad del programa de educación superior en el trabajo

La mayor proporción de graduados consideran que el programa de educación superior ha sido muy útil en el desempeño laboral (67.4%). Pero se observa un grupo de graduados representativo que califica la utilidad del programa de educación en el trabajo como medianamente útil (29,2%) (Gráfica 3).

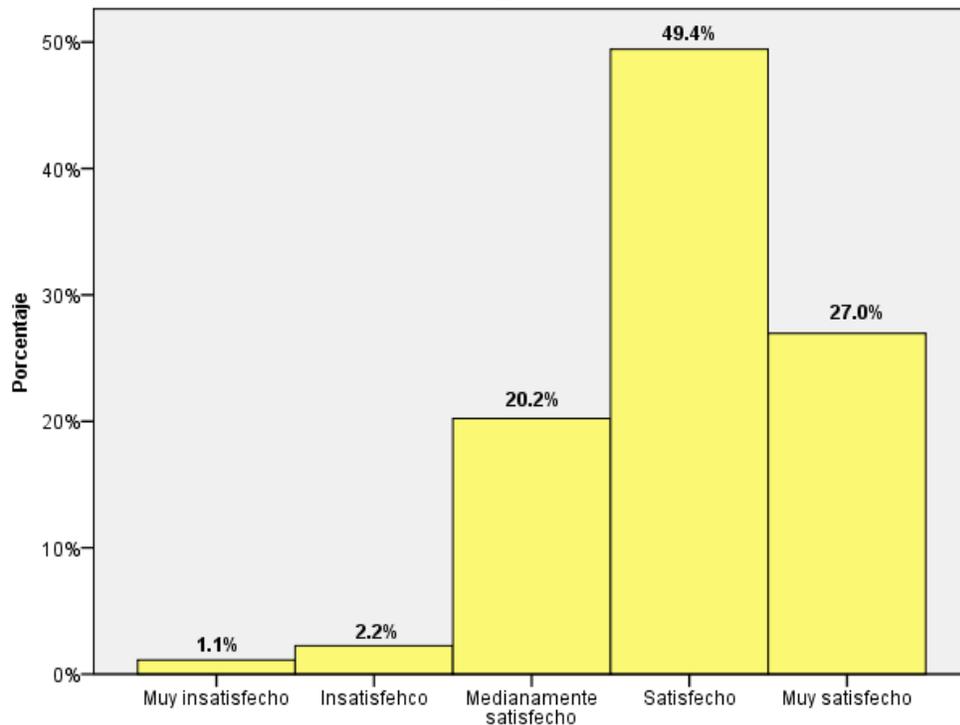
Gráfica 3. Utilidad del programa de educación superior en el trabajo



A4. Satisfacción con el trabajo

Se observa tres calificaciones de los graduados referentes a la satisfacción con el trabajo actual, los satisfechos (49.4%), muy satisfechos (27%) y medianamente satisfechos (20.2%) (Gráfica 4). De manera general, una buena proporción de graduados consideran que su trabajo actual es satisfactorio.

Gráfica 4. Satisfacción con el trabajo actual

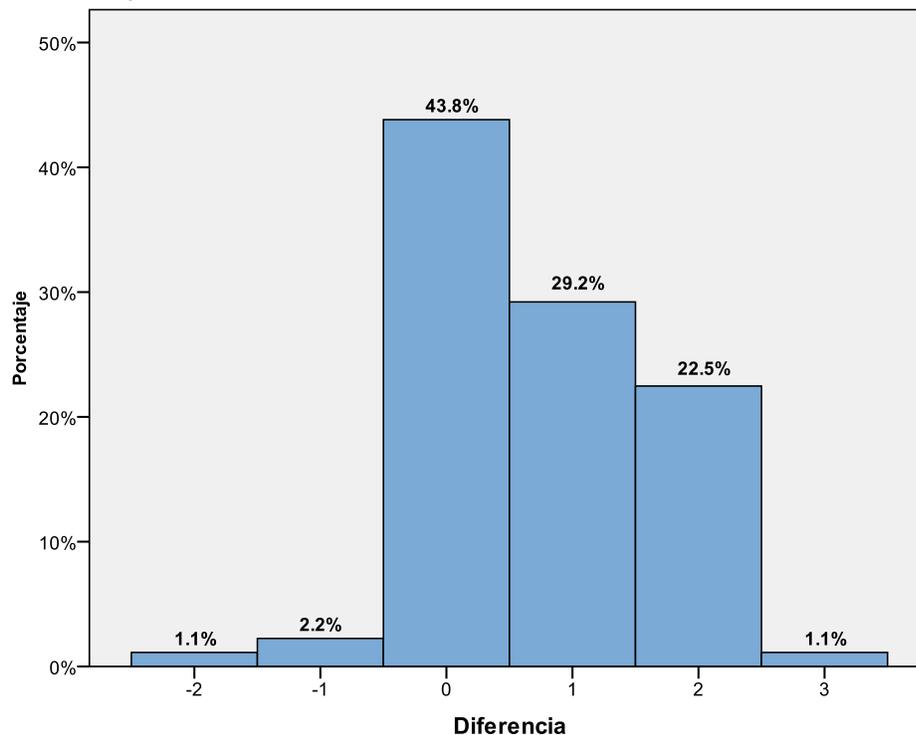


A5. Diferencia entre Importancia dada por el graduado y la empresa a la estabilidad laboral¹⁴

Las diferencias que prevalecen respecto a la estabilidad laboral son la no existencia de diferencias (43.8%), la diferencia entre la importancia dada por el graduado y la empresa son más estrechas, teniendo en cuenta que es 1 (29.2%) y hay una diferencia más relevante entre la expectativa del graduado y la importancia otorgada por la empresa, dado que es 2 (22.5%) (Gráfica 5).

¹⁴ El nivel de importancia va del 1 al 5, donde 1= Nada importante y 5= Muy importante. Las mismas escalas se conservan para la evaluación realizada en referencia a la importancia dada por el graduado y la importancia dada por la empresa. Las diferencias posibles son -4,-3,-2,-1,0,1,2,3,4. Valores negativos indican que la importancia dada por la empresa prevalece sobre la del graduado, valores positivos tienen el significado inverso y cero no hay diferencias. Esta misma escala se conserva para las preguntas desde la A6 a la A8.

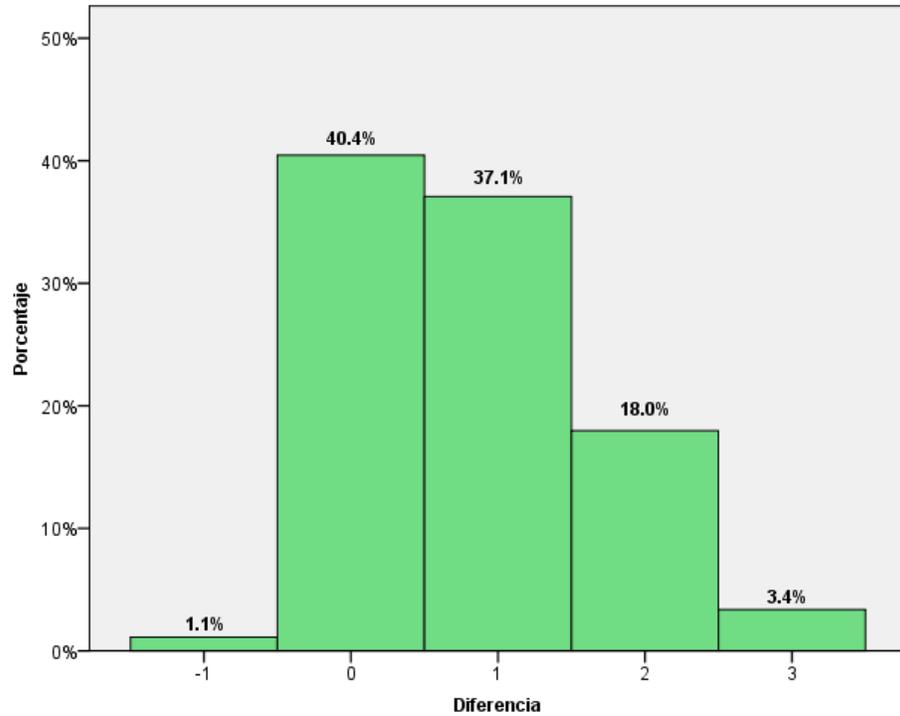
Gráfica 5. Diferencia entre Importancia dada por el graduado y la empresa a la estabilidad laboral



A6. Diferencia entre la importancia dada por el graduado y la empresa a la oportunidad de adquirir nuevos conocimientos.

Las diferencias que prevalecen respecto a la oportunidad de adquirir nuevos conocimientos son la no existencia de diferencias (40.4%), la diferencia entre la importancia dada por el graduado y la empresa son más estrechas, teniendo en cuenta que es 1 (37.1%) y hay una diferencia más relevante entre la expectativa del graduado y la importancia otorgada por la empresa, dado que es 2 (18%) (Gráfica 5).

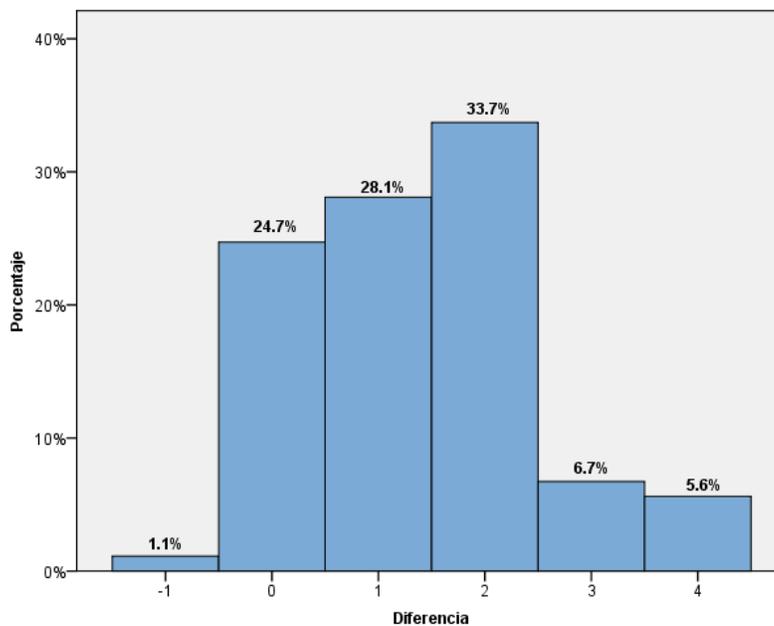
Gráfica 6. Diferencia entre Importancia dada por el graduado y la empresa a la oportunidad de adquirir nuevos conocimientos



A7. Diferencia entre la importancia dada por el graduado y la empresa a la mejora de los ingresos laborales.

La relación entre la importancia dada por el graduado y la empresa a la mejora de los ingresos laborales es considerable, dado que es dos. Prevalciendo la importancia dada por el graduado (33.7%), van disminuyendo a 1 (28.1%) y luego a cero (24.7%) (Gráfica 7).

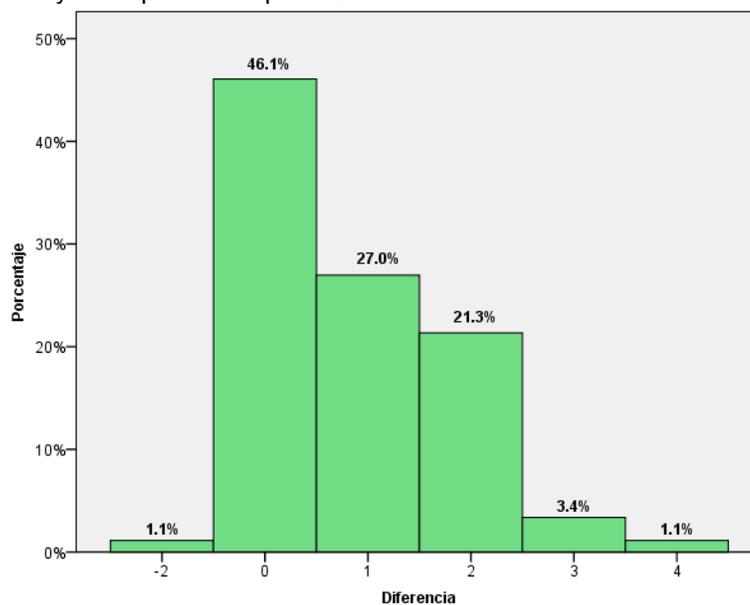
Gráfica 7. Diferencia entre Importancia dada por el graduado y la empresa a la mejora de los ingresos laborales



A8. Diferencia entre la importancia dada por el graduado y la empresa a la posibilidad de afrontar nuevos retos.

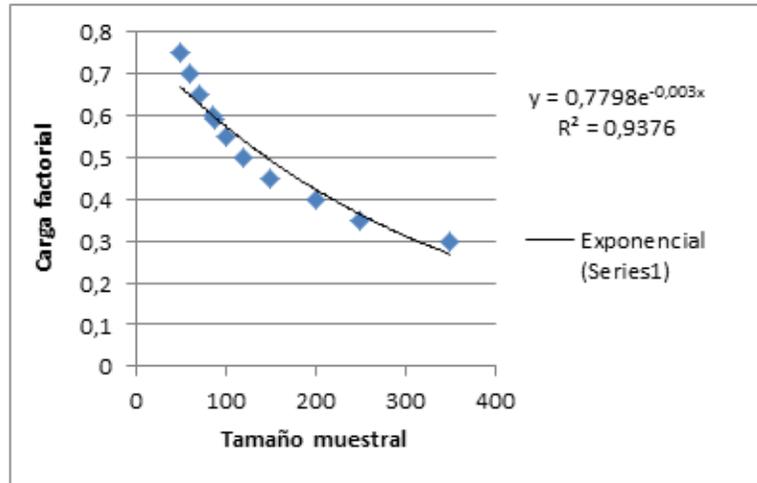
La diferencia que prevalece respecto a la posibilidad de afrontar nuevos retos, es la no existencia de diferencias (46.1%) (Gráfica 8)

Gráfica 8. Diferencia entre Importancia dada por el graduado y la empresa a la posibilidad de afrontar nuevos retos



ANEXO B. Directrices para la identificación de las cargas factoriales significativas basadas en el tamaño muestral, con nivel de significación del 0,05 y un nivel del potencia del 80%

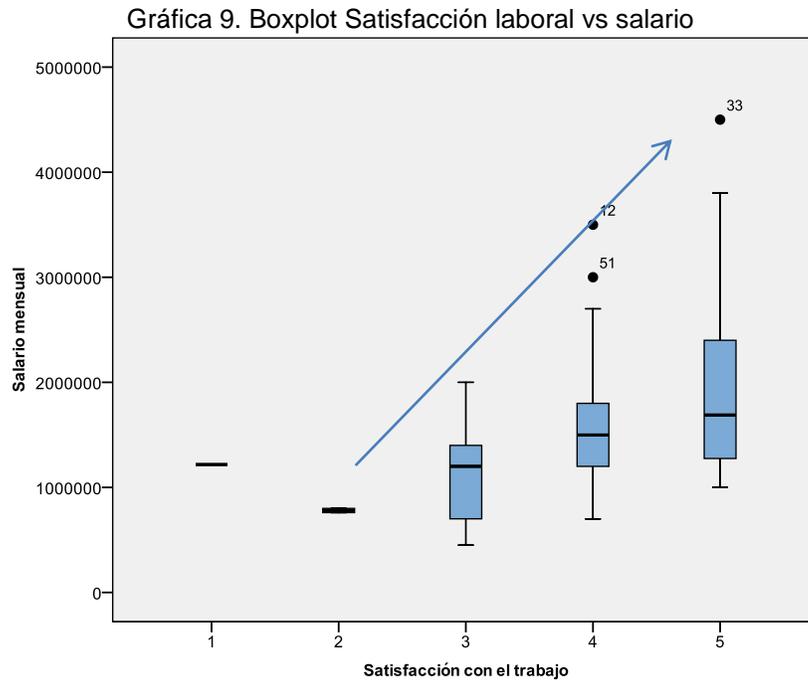
TAMAÑO MUESTRAL	CARGA FACTORIAL
350	0,3
250	0,35
200	0,4
150	0,45
120	0,5
100	0,55
89	0,589
85	0,6
70	0,65
60	0,7
50	0,75



ANEXO C. Análisis descriptivo bivariado

B1. Satisfacción laboral vs salario

Se describe una tendencia positiva, es decir a mayor salario mayor satisfacción laboral (Gráfica 9).



B2. Relación del trabajo con el programa de educación superior vs utilidad del programa de educación superior en el trabajo

Hay una tendencia positiva, es decir, a mayor relación del trabajo con el programa de educación superior mayor utilidad de los conocimientos y habilidades aprendidas en la actividad laboral (Cuadro 9).

Cuadro 9. Tabla de contingencia A2 vs A3

A3. Utilidad del programa de educación superior en el trabajo	A2. Relación del trabajo con el programa de educación superior culminado				Total
	1	2	3	4	
1				1,5%	1,1%
2	50,0%		6,3%		2,2%
3	50,0%	33,3%	56,3%	22,1%	29,2%
4		66,7%	37,5%	76,5%	67,4%
Total	100,0%	100,0%	100,0%	100,0%	100,0%

B3. Diferencia entre la importancia dada por el graduado y la empresa a la posibilidad de afrontar nuevos retos vs la diferencia entre la importancia dada por el graduado y la empresa a la mejora de los ingresos laborales.

Hay una tendencia positiva, entre estas dos variables (Cuadro 10).

Cuadro 10. Tabla de contingencia A2 vs A3

A7. Diferencia entre la importancia dada por el graduado y la empresa a la mejora de los ingresos laborales	A8. Diferencia entre la importancia dada por el graduado y la empresa a la posibilidad de afrontar nuevos retos						Total
	-2	0	1	2	3	4	
-1		2,4%					1,1%
0	100,0%	41,5%	8,3%	10,5%			24,7%
1		19,5%	50,0%	21,1%	33,3%		28,1%
2		31,7%	33,3%	47,4%			33,7%
3		2,4%	4,2%	10,5%	66,7%		6,7%
4		2,4%	4,2%	10,5%		100,0%	5,6%
Total	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%