

Trabajo encargado de métodos multivariados

Indicaciones

- El trabajo se puede realizar en colaboración de 4 participantes como máximo.
- Se debe cargar al aula virtual, el script de R y un informe con los resultados.
- Plazo de entrega 07 de abril hasta 11:59 pm.

Descripción del conjunto de datos

La base de datos, que consta de 200 observaciones sobre 23 variables separadas, se basa en un mercado estudio de segmentación de clientes HBAT. En las encuestas se recogieron dos tipos de información. El primer tipo de información fue percepciones del desempeño de HBAT en 13 atributos en la que se consideran las más influyentes en la selección de proveedores en la industria papelera. Los encuestados incluyeron gerentes de compras de firmas que compran a HBAT, y calificaron a HBAT en cada uno de los 13 atributos utilizando una escala de 0 a 10, siendo 10 “Excelente” y 0 “Deficiente”. Dichas funciones comerciales se midieron en una escala de calificación gráfica, donde se dibujó una línea de 10 centímetros entre los puntos finales, etiquetada “Deficiente y excelente”. Como parte de la encuesta, los encuestados indicaron sus percepciones haciendo una marca en cualquier parte de la línea. Se observó electrónicamente la ubicación de la marca y la distancia desde 0 (en centímetros) se registró en la base de datos para esa encuesta en particular. Los resultados fueron una escala de 0 a 10, redondeado a un solo decimal. El segundo tipo de información se relaciona con resultados de compra y relaciones comerciales (p. ej., satisfacción con HBAT y si la empresa considerará una alianza/asociación estratégica con HBAT).

Variables demográficas de los clientes

Variables	Descripción
X1 Tipo de cliente	Tiempo que un cliente en particular ha estado comprando en HBAT: 1= Menos de un año, 2= Entre 1 y 5 años, 3= Más de 5 años
X2 Tipo de industria	Tipo de industria que compra los productos de papel de HBAT: 0= Industria de revistas, 1= industria del papel prensa
X3 Tamaño de empresa	Tamaño de la empresa Tamaño del empleado: 0= Empresa pequeña, menos de 500 empleados, 1= Empresa grande, 500 o más empleados
X4 Región	Ubicación del cliente de la región X4: 0= EE. UU./América del Norte, 1= Fuera de Norteamérica
X5 Sistema de distribución.	Cómo se venden los productos de papel a los clientes: 0= Vendido indirectamente a través de un corredor, 1= Vendido directamente

Variables de percepción del desempeño

Variables	Descripción
X6 Calidad del producto	Nivel percibido de calidad de los productos de papel de HBAT

X7 Sitio web	Imagen general del sitio web de HBAT, especialmente la facilidad de uso
X8 Soporte técnico	Grado en el que se ofrece soporte técnico para ayudar resolver problemas de productos/servicios
X9 Resolución de quejas	Grado en que las quejas se resuelven en una manera oportuna y completa
X10 Publicidad	Percepciones de las campañas publicitarias de HBAT en todos tipos de medios
X11 Línea de productos	Profundidad y extensión de la línea de productos de HBAT para satisfacer las necesidades del consumidor
X12 Imagen de la fuerza de ventas	Imagen general de la fuerza de ventas de HBAT
X13 Precios competitivos	Grado en el que HBAT ofrece precios competitivos
X14 Garantía y reclamaciones	Grado en que HBAT respalda su producto/ servicio garantías y reclamos
X15 Nuevos productos	Grado en que HBAT desarrolla y vende nuevos productos
X16 Pedidos y facturación	Percepciones de que se manejan pedidos y facturación eficiente y correctamente
X17 Flexibilidad de precios	Disposición percibida de los representantes de ventas de HBAT para negociar precio en compras de productos de papel
X18 Velocidad de entrega	Cantidad de tiempo que lleva entregar los productos de papel una vez que se ha confirmado un pedido

Variables de relación con la empresa

Variables	Descripción
X19 Métrica de satisfacción	Satisfacción del cliente con compras anteriores de HABA, medido en una escala de calificación gráfica de 10 puntos
X20 Métrica de probabilidad de recomendación	Probabilidad de recomendar HABA a otras empresas como proveedor de productos de papel, medido en una escala de 10 puntos
X21 Métrica de probabilidad de compra futura	Probabilidad de comprar productos de papel de HABA en el futuro, medido en una escala de calificación gráfica de 10 puntos
X22 Métrica de nivel de uso/compra actual	Porcentaje de las necesidades de papel de la empresa que respondió comprado de HABA, medido en un punto de 100 escala de porcentaje
X23 Considerar Alianza Estratégica/Asociación en el Futuro No métrico	Percepción de relación futura con HABA o Grado en que el cliente/encuestado percibe su o su empresa participaría en una alianza estratégica /asociación con HABA: 0= No consideraría, 1= Sí, consideraría una alianza o asociación estratégica

Cuestiones por responder

El conjunto de datos se encuentra disponible en el archivo HBAT.xlsx. Con dicha información plantee las hipótesis estadísticas, elija el método estadístico adecuado y aplíquelo para desarrollar 3 de las siguientes cuestiones:

1. ¿Se cree que las empresas que van comprando más de 5 años presentan mejores métricas para las variables de relación con la empresa?
2. ¿Es posible encontrar factores o componentes que reduzcan la dimensionalidad de las Variables de percepción del desempeño? ¿cómo los nombraría?
3. ¿Es posible encontrar factores o componentes que reduzcan la dimensionalidad de las Variables demográficas de los clientes y variables de percepción del desempeño? ¿cómo los nombraría?
4. ¿Es posible encontrar segmentos de clientes? ¿Cuántos segmentos son? ¿Qué características poseen dichos segmentos?
5. ¿Considerando las variables percepción del desempeño o los factores que reducen su dimensionalidad, estime un modelo de regresión logística o árboles de clasificación para predecir si el cliente realizará una alianza estratégica con HBAT?
6. Categoriza a los clientes de Métrica de satisfacción como satisfechos para un puntaje mayor a 7, insatisfechos para un puntaje menor a 4 y el resto como parcialmente satisfechos. ¿Considerando las variables percepción del desempeño o los factores que reducen su dimensionalidad, estime un modelo de regresión logística ordinal para predecir el nivel de satisfacción de los clientes de la empresa HBAT?

1. **¿Se cree que las empresas que van comprando más de 5 años presentan mejores métricas para las variables de relación con la empresa?**

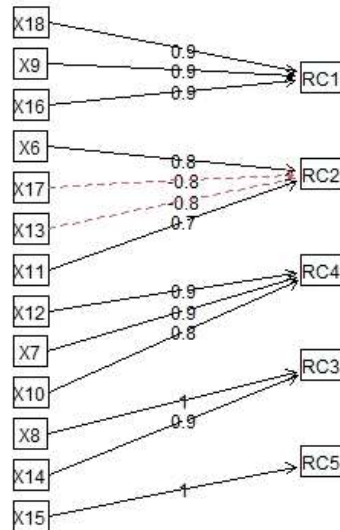
Al revisar las medias vemos que así es el grupo “A” está definido por las que tienen Menos de un año, el grupo “B” esta entre 1 y 5 años, y el grupo C con más de 5 años. Ya que al ordenarlo de mayor a menor vemos que el grupo C es el que presenta mejores métricas en algunos puntos solo siendo superado en muy pocos por los otros grupos.

	g	variable	n	mean	sd
	<fct>	<fct>	<dbl>	<dbl>	<dbl>
1	A	X19	68	5.73	0.764
2	B	X19	64	7.29	0.708
3	C	X19	68	7.85	1.03
4	A	X20	68	6.14	0.995
5	B	X20	64	7.21	0.714
6	C	X20	68	7.52	0.976
7	A	X21	68	6.96	0.76
8	B	X21	64	7.88	0.643
9	C	X21	68	8.16	0.777
10	A	X22	68	49.0	5.02
11	B	X22	64	58.6	5.26
12	C	X22	68	67.0	4.84
13	A	X23	68	0.074	0.263
14	B	X23	64	0.531	0.503
15	C	X23	68	0.691	0.465

2. ¿Es posible encontrar factores o componentes que reduzcan la dimensionalidad de las Variables de percepción del desempeño? ¿cómo los nombraría?

Usando análisis de componentes principales si es posible encontrar factores que reduzcan la dimensionalidad donde encontramos 4 a los cuales describimos abajo y nombramos

Components Analysis



X18 Velocidad de entrega
X9 Resolución de quejas
X16 Pedidos y facturación

Atención al cliente

X6 Calidad del producto
X17 Flexibilidad de precios
X13 Precios competitivos
X11 Línea de productos

Competitividad

X12 Imagen de la fuerza de ventas
X7 Sitio web
X10 Publicidad

Marketing

X8 Soporte técnico
X14 Garantía y reclamaciones

Asistencia al cliente

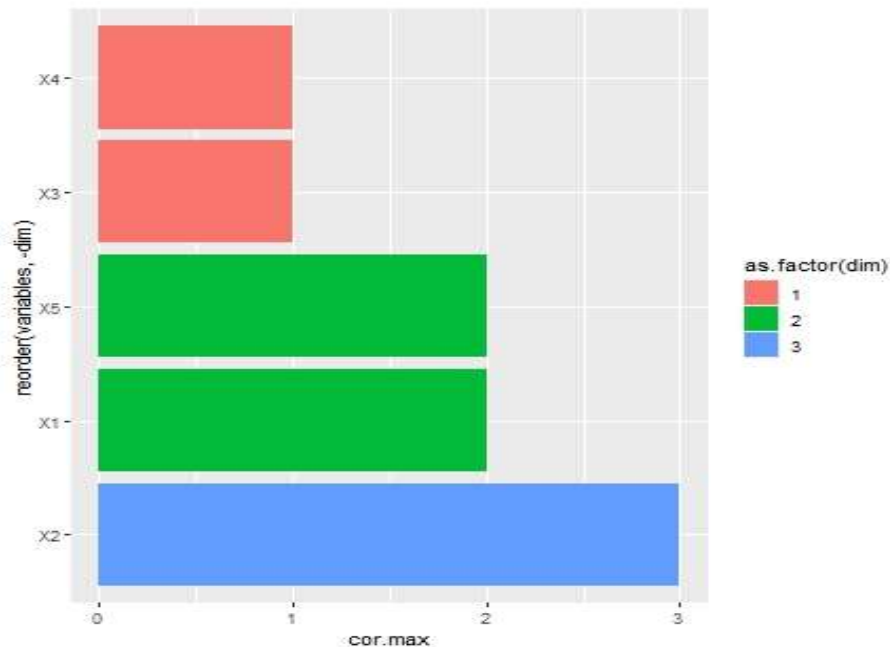
X15 Nuevos productos

Nuevos productos

3. ¿Es posible encontrar factores o componentes que reduzcan la dimensionalidad de las Variables demográficas de los clientes y variables de percepción del desempeño?
¿cómo los nombraría?

Para Variables demográficas de los clientes

Podríamos identificar 3 componentes



X3 Tamaño de empresa

X4 Región

Ubicación clientes

X1 Tipo de cliente

X5 Sistema de distribución.

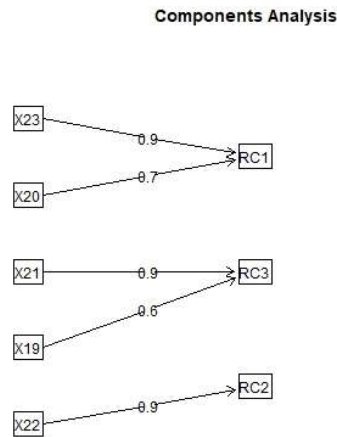
Embarques

X2 Tipo de industria

Industria

Para Variables de relación con la empresa

Hay una alta correlación entre la información, pero podríamos hacer una separación en 3 componente para reducir el análisis



X23 Considerar Alianza Estratégica/Asociación en el Futuro No métrico
X20 Métrica de probabilidad de recomendación

Alianzas Empresariales

X21 Métrica de probabilidad de compra futura
X19 Métrica de satisfacción

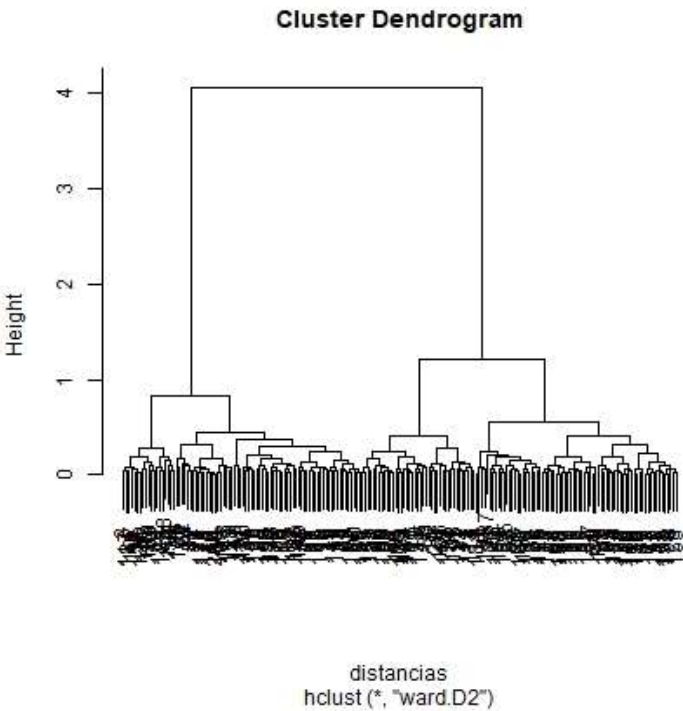
Satisfacción de la empresa

X22 Métrica de nivel de uso/compra actual

Compras actuales

4. ¿Es posible encontrar segmentos de clientes? ¿Cuántos segmentos son? ¿Qué características poseen dichos segmentos?

Si se puede segmentar en 3



5. ¿Considerando las variables percepción del desempeño o los factores que reducen su dimensionalidad, estime un modelo de regresión logística o arboles de clasificación para predecir si el cliente realizará una alianza estratégica con HBAT?

El Modelo tendría los siguientes coeficientes

Coefficients:		
	Values	Std. Err.
(Intercept)	-28.611752652	5.6730289
X6	1.088959071	0.2173522
X7	0.294289791	0.4693111
X8	-0.068418684	0.2279858
X9	-0.026080769	0.3687155
X10	-0.435672749	0.2448456
X11	2.416075772	1.2274747
X12	1.146606945	0.3812065
X13	-0.355008230	0.1673602
X14	0.176390712	0.4355283
X15	-0.001193901	0.1439745
X16	-0.039339008	0.3550625
X17	2.858528239	1.2608088
X18	-2.818907924	2.3418845
Residual Deviance: 157.0433		
AIC: 185.0433		

6. Categoriza a los clientes de Métrica de satisfacción como satisfechos para un puntaje mayor a 7, insatisfechos para un puntaje menor a 4 y el resto como parcialmente satisfechos. ¿Considerando las variables percepción del desempeño o los factores que reducen su dimensionalidad, estime un modelo de regresión logística ordinal para predecir el nivel de satisfacción de los clientes de la empresa HBAT?

Los coeficientes el modelo de regresión logística serian:

Coefficients:		
	Values	Std. Err.
(Intercept)	-0.31212982	2.8834685
X6	0.03570402	0.1394008
X7	-0.08324524	0.3145604
X8	0.19775184	0.1680755
X9	-0.15863566	0.2665273
X10	0.02207304	0.1743410
X11	0.02030339	0.6719469
X12	-0.21636858	0.2507387
X13	0.02595496	0.1180186
X14	-0.31584362	0.3274035
X15	-0.04100043	0.1027084
X16	-0.14173597	0.2692489
X17	0.30915147	0.7004419
X18	0.63499522	1.3585825
Residual Deviance: 261.9576		
AIC: 289.9576		

Para la regresión logística ordinal se necesitan al menos 3 niveles, pero con la categorización propuesta solo se tienen 2 ya que no hay datos con puntajes menores de 4, por lo que no se podría obtener.