**Escuela de Estudios de Postgrado**

**Facultad de Ingeniería**

**Universidad San Carlos de Guatemala**

**Maestría en Ingeniería para la Industria con**

**Especialización en Ciencias de la Computación.**

**Minería de Datos**

**Tarea 1**

**ALGORITMO APRIORI**

**Néstor Obdulio González López**

**Carné: 999013150**

Descripción de 3 reglas de asociación

persona <- read.csv("C:\\Users\\nestor.gonzalez\\Desktop\\Master Gadex\\MAESTRIA USAC\\INTRODUCCION A LA MINERIA DE DATOS\\sabado\_5\_10\\PERSONA\_BDP.csv")

unique(persona$PEA)

persona$PEI[is.na(persona$PEI)] <- 0

persona <- persona[, sapply(persona, function(x) length(unique(x)) > 1)]

persona$PEI <- factor(persona$PEI, levels = c(0, 1), labels = c("No", "Yes"))

persona$PEI <- as.numeric(as.character(persona$PEI))

persona$PEI <- cut(persona$PEI, breaks = c(-Inf, 0, Inf), labels = c("Bajo", "Alto"))

library(arules)

reglas <- apriori(persona, parameter = list(support = 0.2, confidence = 0.5))

inspect(reglas[0:500])

PCP15[1,25] => PCP12[1,5)

SUPPORT = 0.26

CONFIDENCE = 0.98

COVERAGE = 0.27

|  |  |  |
| --- | --- | --- |
| PCP15 | 25 | ¿Cuál es el idioma en el que aprendió a hablar? |
| PCP15 | 1 | Achí |
|  | 2 | Akateko |
|  | 3 | Awakateko |
|  | 4 | Ch'orti' |
|  | 5 | Chalchiteko |
|  | 6 | Chuj |
|  | 7 | Itza' |
|  | 8 | Ixil |
|  | 9 | Jakalteko/Popti' |
|  | 10 | K'iche' |
|  | 11 | Kaqchiquel |
|  | 12 | Mam |
|  | 13 | Mopan |
|  | 14 | Poqomam |
|  | 15 | Poqomchi' |
|  | 16 | Q'anjob'al |
|  | 17 | Q'eqchi' |
|  | 18 | Sakapulteko |
|  | 19 | Sipakapense |
|  | 20 | Tektiteko |
|  | 21 | Tz'utujil |
|  | 22 | Uspanteko |
|  | 23 | Xinka |
|  | 24 | Garífuna |
|  | 25 | Español |
|  | 26 | Inglés |
|  | 27 | Señas |
|  | 28 | Otro idioma |
|  | 98 | No habla |

|  |  |  |
| --- | --- | --- |
| PCP12 | 22 | Según su origen o historia,¿cómo se considera o auto identifica? |
| PCP12 | 1 | Maya |
|  | 2 | Garífuna |
|  | 3 | Xinka |
|  | 4 | Afrodescendiente/Creole/Afromestizo |
|  | 5 | Ladina (o) |
|  | 6 | Extranjera (o) |

La regla de asociación refleja un patrón lógico y esperado en el censo, en el que las personas que hablan un idioma indígena o minoritario tienden a identificarse culturalmente con su grupo étnico. Esto se debe a la preservación de la lengua como un elemento central de la identidad cultural y al reconocimiento de las raíces y tradiciones que estas lenguas representan.

PCP11\_B[1,6) => RESCINGEO[101,601)

SUPPORT = 0.29

CONFIDENCE = 1.0

COVERAGE = 0.29

|  |  |  |
| --- | --- | --- |
| PCP11\_B | 19 | Departamento de residencia en abril de 2013 |

|  |  |
| --- | --- |
| 1 | Guatemala |
| 2 | El Progreso |
| 3 | Sacatepéquez |
| 4 | Chimaltenango |
| 5 | Escuintla |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RESCINGEO | | 20 | Municipio de residencia en abril de 2013 | |
| 101 | Guatemala | | |
| 102 | Santa Catarina Pinula | | |
| 103 | San José Pinula | | |
| 104 | San José del Golfo | | |
| 105 | Palencia | | |
| 106 | Chinautla | | |
| 107 | San Pedro Ayampuc | | |
| 108 | Mixco | | |
| 109 | San Pedro Sacatepéquez | | |
| 110 | San Juan Sacatepéquez | | |
| 111 | San Raymundo | | |
| 112 | Chuarrancho | | |
| 113 | Fraijanes | | |
| 114 | Amatitlán | | |
| 115 | Villa Nueva | | |
| 116 | Villa Canales | | |
| 117 | Petapa | | |
| 199 | Guatemala municipio no declarado | | |
| 201 | Guastatoya | | |
| 202 | Morazán | | |
| 203 | San Agustín Acasaguastlán | | |
| 204 | San Cristóbal Acasaguastlán | | |
| 205 | El Jícaro | | |
| 206 | Sansare | | |
| 207 | Sanarate | | |
| 208 | San Antonio la Paz | | |
| 299 | El Progreso municipio no declarado | | |
| 301 | Antigua Guatemala | | |
| 302 | Jocotenango | | |
| 303 | Pastores | | |
| 304 | Sumpango | | |
| 305 | Santo Domingo Xenacoj | | |
| 306 | Santiago Sacatepéquez | | |
| 307 | San Bartolomé Milpas Altas | | |
| 308 | San Lucas Sacatepéquez | | |
| 309 | Santa Lucía Milpas Altas | | |
| 310 | Magdalena Milpas Altas | | |
| 311 | Santa María de Jesús | | |
| 312 | Ciudad Vieja | | |
| 313 | San Miguel Dueñas | | |
| 314 | San Juan Alotenango | | |
| 315 | San Antonio Aguas Calientes | | |
| 316 | Santa Catarina Barahona | | |
| 399 | Sacatepéquez municipio no declarado | | |
| 401 | Chimaltenango | | |
| 402 | San José Poaquil | | |
| 403 | San Martín Jilotepeque | | |
| 404 | Comalapa | | |
| 405 | Santa Apolonia | | |
| 406 | Tecpán Guatemala | | |
| 407 | Patzún | | |
| 408 | San Miguel Pochuta | | |
| 409 | Patzicía | | |
| 410 | Santa Cruz Balanyá | | |
| 411 | Acatenango | | |
| 412 | Yepocapa | | |
| 413 | San Andrés Itzapa | | |
| 414 | Parramos | | |
| 415 | Zaragoza | | |
| 416 | El Tejar | | |
| 499 | Chimaltenango municipio no declarado | | |
| 501 | Escuintla | | |
| 502 | Santa Lucía Cotzumalguapa | | |
| 503 | La Democracia | | |
| 504 | Siquinalá | | |
| 505 | Masagua | | |
| 506 | Tiquisate | | |
| 507 | La Gomera | | |
| 508 | Guanagazapa | | |
| 509 | San José | | |
| 510 | Iztapa | | |
| 511 | Palín | | |
| 512 | San Vicente Pacaya | | |
| 513 | Nueva Concepción | | |
| 514 | Sipacate | | |
| 599 | Escuintla municipio no declarado | | |

La regla de asociación muestra una relación esperada y lógica entre el departamento y el municipio de residencia. La confianza del 100% y el lift elevado confirman que las divisiones administrativas están bien representadas y codificadas en el censo, lo cual es consistente con la estructura geográfica y administrativa de Guatemala

COD\_MUNICIPIO[1,4) => PCP12[5,6)

SUPPORT = 0.20

CONFIDENCE = 0.67

COVERAGE = 0.30