

Análisis Multivariado
Taller No. 5: ACS de la TC manzanas de Bogotá según localidades
y estratos

Laura Elizabeth Holguín
Eduardo Contreras Bohórquez
Diana Catalina Martínez

14 de septiembre de 2017

1. Comente la repartición de las manzanas según estratos (histograma de la distribución de las manzanas en los 6 estratos - distribución marginal)

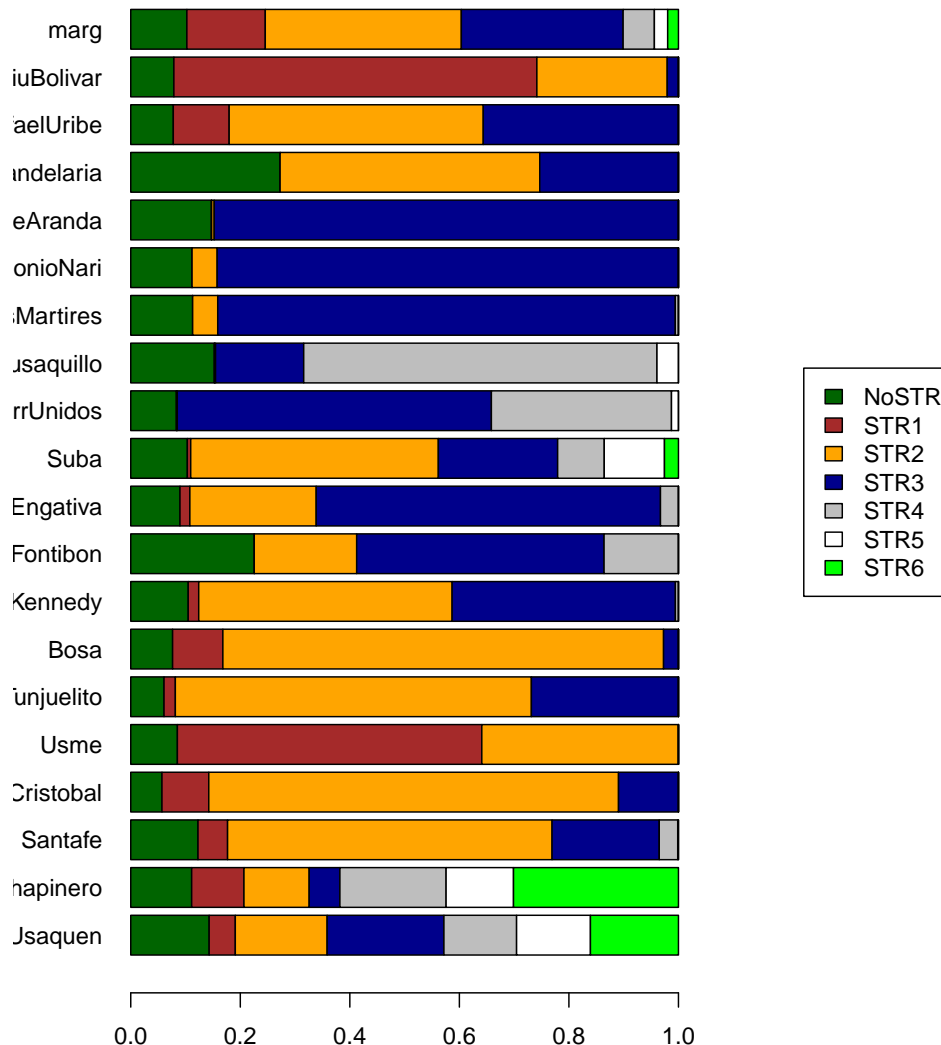


Figura 1: Perfiles de las manzanas en Bogotá según estratos

Según el histograma de perfiles por estrato (figura 1), en la distribución marginal los estratos 2 y 3 son los que más manzanas tienen, es decir, geográficamente Bogotá tiende a concentrar habitantes de clase media-baja. En contraste, los estratos más altos, 4, 5 y 6, son los que están en menor proporción de manzanas. Las manzanas no estratificadas, por su parte, corresponden aproximadamente al 10% del total de la información. El estrato con menor proporción de manzanas es el 5.

2. ¿Cómo es la distribución de las manzanas según localidades (distribución marginal)?

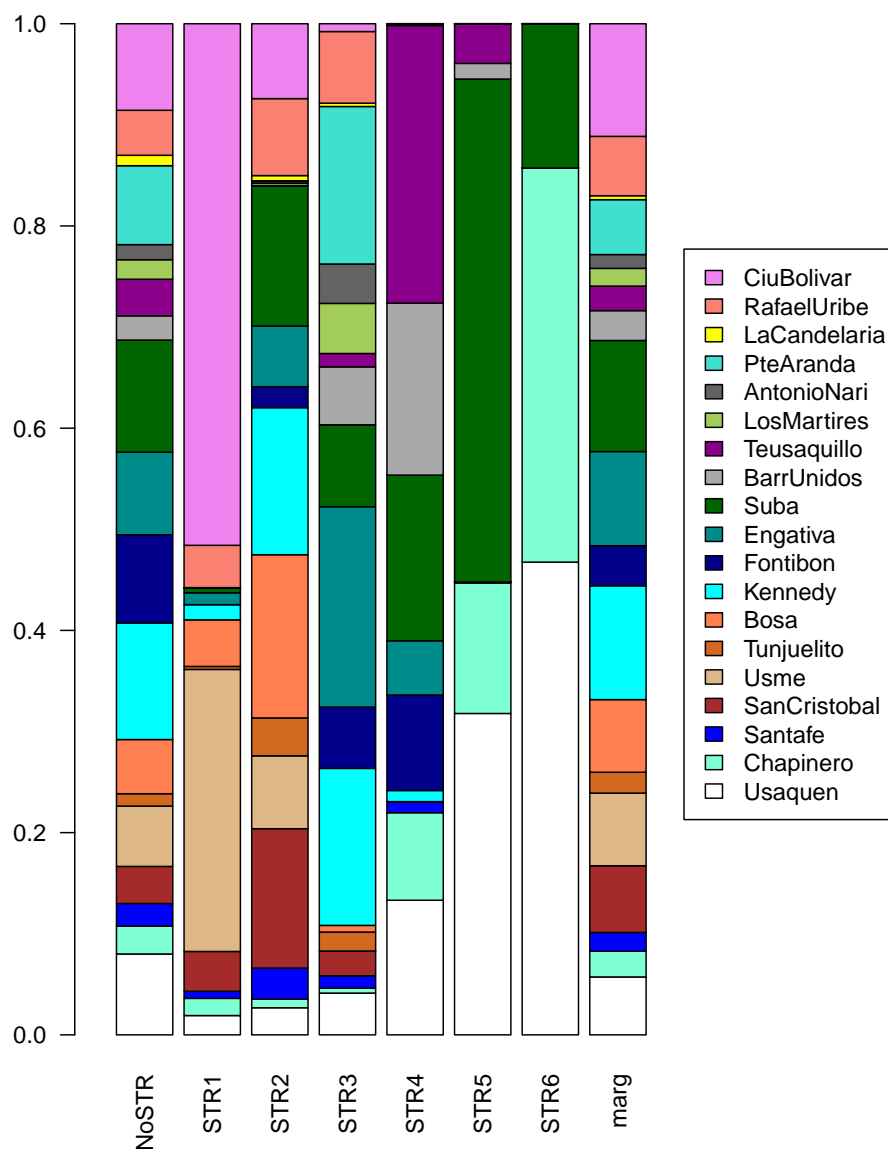


Figura 2: Perfiles de la distribución de las manzanas en Bogotá según localidades

Las localidades con mayor proporción de manzanas, de acuerdo al histograma (figura 2), son Ciudad Bolívar, Suba y Kennedy, respectivamente. En contraste, la Candelaria, Antonio Nariño y Los Mártires son las localidades con menor proporción de manzanas. En general, ninguna localidad sobrepasa el 12 % de dicha proporción.

3. ¿Utilizaría la columna sin estrato como activa en un análisis de correspondencias?; ¿Porque sí? ; ¿Por qué no?

La columna sin estrato corresponde a manzanas que contienen áreas comunes, servicios sociales, entre otros. Aunque dichas zonas podrían estar relacionadas con la estratificación, estas no son relevantes para conocer cómo es la distribución de las localidades y los estratos, así mismo, la distribución de las manzanas no estratificadas no tiene una tendencia clara en las localidades, es decir, ni por perfiles fila ni por columna hay un patrón de dominancia. Por tanto, esta columna no sería utilizada como activa en un análisis de correspondencias, sino como variable ilustrativa.

4. Compare la estadística chi2 asociada a la tabla de contingencia con la teórica. ¿Hay asociación entre estratos y localidades?

Se realizó una prueba de Independencia χ^2 entre las variables estrato y localidad para los datos reportados en "Bogotá", la cual arrojó evidencias de que dichas variables no se comportan de manera independiente, rechazando H_0 con un p valor de 2.2 exp -16 y un nivel de confianza del 95 %. De igual forma se realizó la comparación entre los estadísticos teóricos (media y varianza) y el estadístico calculado χ^2 los cuales dieron por resultado 152.1 y 50500.99, respectivamente, con lo cual se concluye que se rechaza H_0 .

Estas variables pueden estar asociadas, pues existen factores como preferencia y capacidad adquisitiva que determinan la demanda de vivienda en las diferentes localidades que se caracterizan por su estrato predominante, así una población con nivel adquisitivo alto demanda vivienda estrato 6 y prefiere que sea en una localidad cuyo estrato dominante es el 6.

	STR1	STR2	STR3	STR4	STR5	STR6
Usaquen	0.00	0.01	0.01	0.01	0.01	0.01
Chapinero	0.00	0.00	0.00	0.01	0.00	0.01
SanCristobal	0.01	0.06	0.01	0.00	0.00	0.00
Usme	0.04	0.03	0.00	0.00	0.00	0.00
Tunjuelito	0.00	0.01	0.01	0.00	0.00	0.00
Bosa	0.01	0.06	0.00	0.00	0.00	0.00
Kennedy	0.00	0.06	0.05	0.00	0.00	0.00
Fontibon	0.00	0.01	0.02	0.01	0.00	0.00
Engativá	0.00	0.02	0.07	0.00	0.00	0.00
Suba	0.00	0.06	0.03	0.01	0.01	0.00
BarrUnidos	0.00	0.00	0.02	0.01	0.00	0.00
Teusaquillo	0.00	0.00	0.00	0.02	0.00	0.00
LosMartires	0.00	0.00	0.02	0.00	0.00	0.00
AntonioNari	0.00	0.00	0.01	0.00	0.00	0.00
PteAranda	0.00	0.00	0.05	0.00	0.00	0.00
LaCandelaria	0.00	0.00	0.00	0.00	0.00	0.00
RafaelUribe	0.01	0.03	0.02	0.00	0.00	0.00
CiuBolivar	0.08	0.03	0.00	0.00	0.00	0.00

	NoSTR	STR1	STR2	STR3	STR4	STR5	STR6
Usaquen	0.05	0.07	0.17	0.14	0.03	0.01	0.01
Chapinero	0.04	0.06	0.14	0.12	0.02	0.01	0.01
SanCristobal	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Usme	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tunjuelito	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bosa	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kennedy	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fontibon	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Engativá	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Suba	0.01	0.02	0.05	0.04	0.01	0.00	0.00
BarrUnidos	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Teusaquillo	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LosMartires	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AntonioNari	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PteAranda	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LaCandelaria	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RafaelUribe	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CiuBolivar	0.00	0.00	0.00	0.00	0.00	0.00	0.00

(a) Tabla de frecuencias **F**

(b) Tabla de Independencia **A**

	NoSTR	STR1	STR2	STR3	STR4	STR5	STR6
Usaquen	0.05	0.07	0.17	0.14	0.03	0.01	0.01
Chapinero	0.04	0.06	0.14	0.12	0.02	0.01	0.01
SanCristobal	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Usme	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tunjuelito	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bosa	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kennedy	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fontibon	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Engativá	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Suba	0.01	0.02	0.05	0.04	0.01	0.00	0.00
BarrUnidos	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Teusaquillo	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LosMartires	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AntonioNari	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PteAranda	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LaCandelaria	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RafaelUribe	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CiuBolivar	0.00	0.00	0.00	0.00	0.00	0.00	0.00

(c) Diferencia **F-A**

Cuadro 1: Tabla de frecuencias relativas de las manzanas de Bogotá (izquierda), de indenpedencia (centro) y diferencia (derecha)

5. ¿Cuántos ejes retiene para el análisis? ¿Por qué?

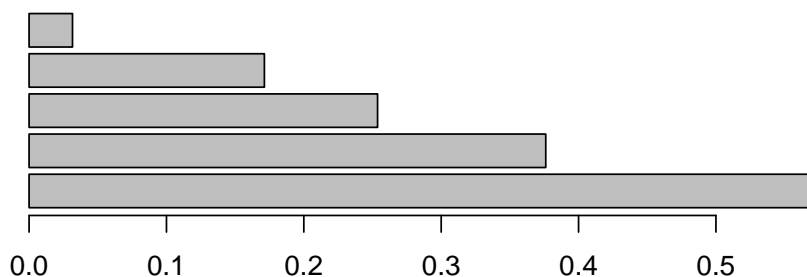


Figura 3: Histograma para los vectores propios

Para el análisis tomaremos los dos primeros ejes, teniendo en cuenta que la inercia recogida por dichos ejes es superior al promedio (0.28), lo cual se puede ver en el respectivo histograma (Figura 3). Los ejes escogidos acumulan el 67 % de la inercia total.

	Coordenadas			Contribuciones Absolutas			Contribuciones Relativas		
	Eje 1	Eje 2	Eje 3	Eje 1(%)	Eje 2(%)	Eje 3(%)	Eje 1	Eje 2	Eje 3
Usaquen	-0.57	-1.16	0.30	3.10	19.54	1.99	-14.73	-61.11	4.19
Chapinero	-0.54	-1.99	0.46	1.29	26.48	2.10	-5.01	-67.81	3.62
Santafe	0.06	0.26	0.44	0.01	0.32	1.39	0.99	19.23	56.86
SanCristobal	0.29	0.29	0.65	1.03	1.59	11.41	12.72	12.90	62.30
Usme	1.26	-0.19	-0.29	20.27	0.68	2.46	92.93	-2.06	-5.00
Tunjuelito	-0.01	0.42	0.49	0.00	1.01	2.07	-0.03	38.76	53.86
Bosa	0.40	0.25	0.78	2.11	1.25	17.83	16.87	6.60	63.27
Kennedy	-0.21	0.47	0.18	0.90	6.58	1.40	-15.14	72.65	10.43
Fontibon	-0.67	0.10	-0.35	2.73	0.08	1.63	-68.66	1.40	-18.22
Engativa	-0.53	0.49	-0.30	4.70	5.92	3.38	-44.86	37.25	-14.34
Suba	-0.31	-0.25	0.46	1.90	1.87	9.16	-17.88	-11.63	38.38
BarrUnidos	-1.04	-0.40	-0.79	5.73	1.29	7.49	-48.86	-7.24	-28.42
Teusaquillo	-1.26	-1.70	-0.82	6.33	17.70	6.07	-18.96	-34.95	-8.08
LosMartires	-0.80	0.68	-0.65	1.94	2.12	2.87	-37.60	27.04	-24.68
AntonioNari	-0.80	0.70	-0.65	1.50	1.75	2.23	-36.50	28.01	-24.04
PteAranda	-0.85	0.72	-0.73	6.49	7.02	10.83	-35.88	25.57	-26.61
LaCandelaria	-0.12	0.46	0.43	0.01	0.18	0.23	-3.17	49.19	42.76
RafaelUribe	0.02	0.38	0.12	0.00	2.36	0.32	0.20	89.36	8.28
CiuBolivar	1.41	-0.27	-0.58	39.96	2.26	15.14	82.63	-3.08	-13.92

Cuadro 2: Coordenadas y ayudas para la interpretación

6. Identifique en el primer eje las localidades más contributivas y sus oposiciones (localidades con coordenadas negativo sobre el eje vs. las de signo positivo).

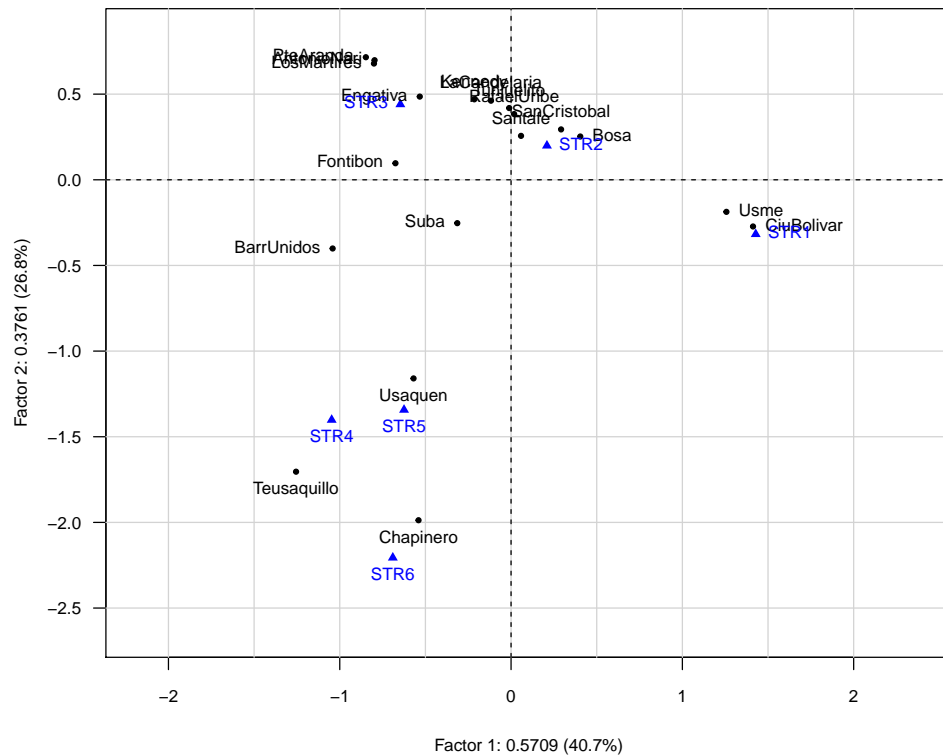


Figura 4: Primer plano factorial para la distribución de las manzanas de Bogotá por localidades

Las localidades que más contribuyen al eje 1 son Ciudad Bolívar y Usme en el eje positivo, a las cuáles se oponen Barrios Unidos, Teusaquillo y Pte Aranda (Figura 4).

7. Identifique los estratos más contributivos al primer eje y sus oposiciones

Los estratos 1, 2 y 3 son los que más contribuyen al primer eje, siendo evidentemente opuestos los estratos 1 y 3, el estrato dos se encuentra representado muy cerca del origen.

8. Repita 6 para el segundo eje

En la figura 4, los estratos que más contribuyen al eje 2 se agrupan en la parte positiva de la recta: Santafé, San Cristóbal, Rafael Uribe, Candelaria, Kennedy, Puente Aranda, Antonio Nariño, Los Mártires y Tunjuelito.

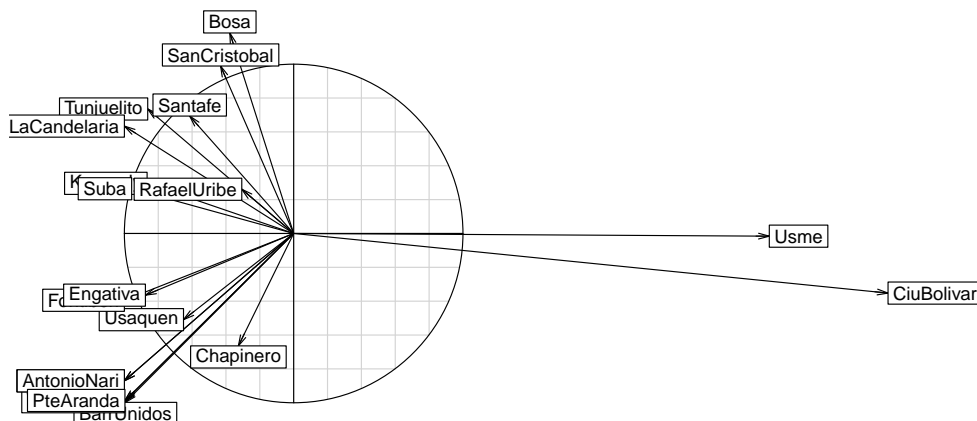


Figura 5: Círculo de correlaciones para el ACS de las manzanas de Bogotá

9. Repita 7 para el segundo eje

Los estratos 1, 2 y 3 se encuentran opuestos a los estratos 4, 5 y 6 según el segundo eje, el que más contribuye es el estrato 2.

10. Resuma la comparación de los perfiles de las localidades utilizando el primer plano factorial.

	DistEuc
Usaquen	1.29
Chapinero	2.06
SantaFe	0.26
SanCristobal	0.41
Usme	1.27
Tunjuelito	0.42
Bosa	0.48
Kennedy	0.52
Fontibon	0.68
Engativa	0.72
Suba	0.40
BarrUnidos	1.12
Teusaquillo	2.12
LosMartires	1.05
AntonioNari	1.06
PteAranda	1.11
LaCandelaria	0.48
RafaelUribe	0.38
CiuBolivar	1.44

Cuadro 3: Distancia Euclidiana para los perfiles por localidades

Los perfiles de las localidades en el primer plano factorial se aglutinan en tres grupos principalmente, la localidad que más se diferencia del promedio es Teusaquillo según cálculos de distancia euclidiana (cuadro 3), las localidades más relacionadas con el primer eje según porcentaje de contribución son Usme y Ciudad Bolívar cuyas cuadras se caracterizan por ser opuestas a las localidades de Fontibón, Barrios Unidos, Puente

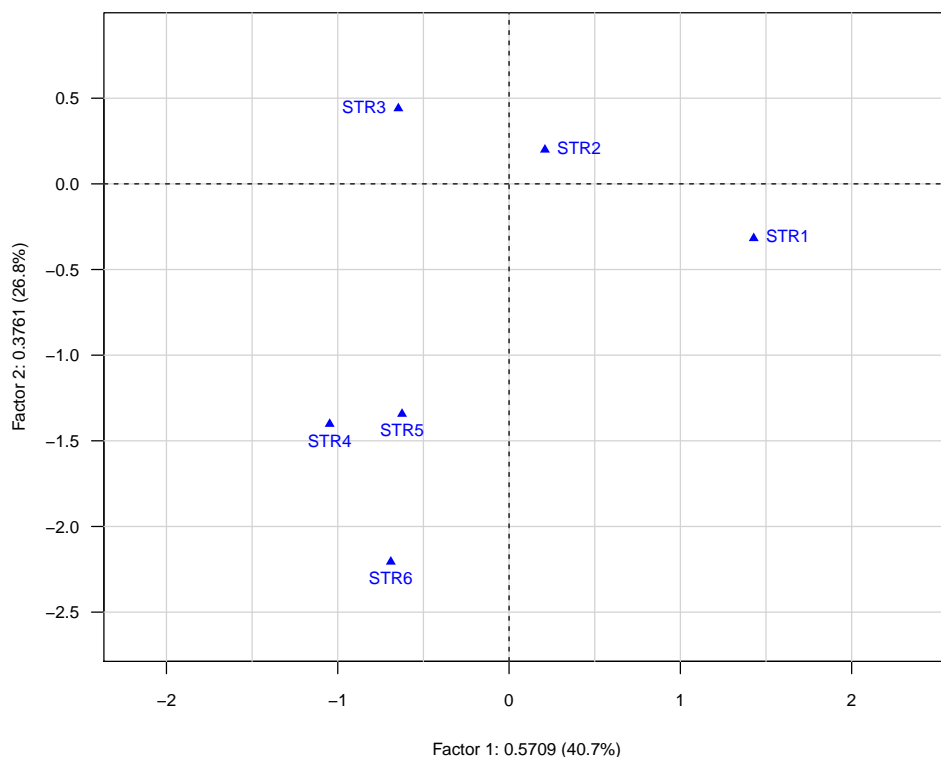


Figura 6: Primer plano factorial para la distribución de las manzanas de Bogotá por estratos

Aranda y Antonio Nariño. La localidad que más se acerca al promedio según cálculos de distancia euclidiana es Santafé. Las localidades que más se relacionan con el segundo eje son Chapinero, Teusaquillo y Usaquén. Las demás localidades tienen manzanas diversas que seguramente contienen diferentes estratos en combinaciones variadas en proporción.

11. Resuma la comparación de los perfiles de los estratos utilizando el primer plano factorial.

Se observa que los perfiles de los estratos se distribuyen ordenadamente desde el estrato 1 al 3 sobre el primer eje en el primer plano factorial, siendo opuestos principalmente estrato 1 y el 4. En el segundo eje se oponen los estratos 3 y 6 y en general el estrato más cercano al promedio es el 2 y el más lejano es el 6 (figura 7).

12. Según el primer plano factorial, ¿cómo es la asociación entre localidades y estratos?

Según el primer plano factorial (figura 4) existen localidades con estratos dominantes que las caracterizan y existen localidades con estratos mixtos. Se pueden observar principalmente tres grupos de localidades definidos por su estrato: estrato 1; estratos 2 y 3; estratos 4, 5 y 6; y unas localidades dispersas entre el estrato 3 y el 4.

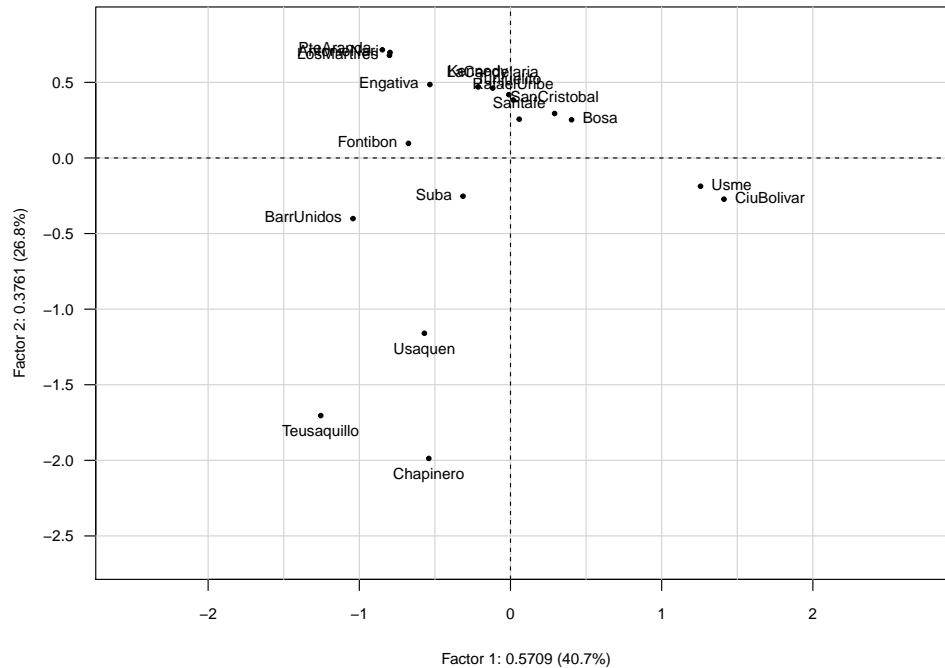


Figura 7: Primer plano factorial para la distribución de las manzanas de Bogotá según localidades

Los estratos más altos agrupan a las localidades de Usaquén, Teusaquillo y Chapinero, los cuales se localizan geográficamente hacia el norte de la ciudad. El estrato 1 se localiza hacia el extremo sur-oriente en las localidades de Usme y Ciudad Bolívar, localidades menos homogéneas en estrato como Barrios Unidos, Fontibón y Suba y la masa de localidades de estrato 2 y 3 se localizan tan aglomeradas que podría decirse que comparten las características de distribución de manzanas en estratos.

13. ¿Hay efecto Guttman? Explique.

De acuerdo al primer plano factorial, se puede observar que se presenta un efecto Guttman. El primer eje opone los estratos extremos, es decir los estratos altos (4, 5 y 6) contra estratos bajos (1 y 2); el eje 2 opone los estratos medios (2 y 3). Así mismo, se observa que los estratos se ordenan siguiendo la parábola y de forma secuencial, de derecha a izquierda aparecen los estratos de menor a mayor. De forma análoga, el orden de los estratos se traslada a las localidades siguiendo una parábola invertida: Usme y Ciudad Bolívar son las localidades que se encuentran en los estratos más bajos, mientras que Usaquén, Teusaquillo y Chapinero están en los estratos más altos.

14. ¿Hay contraposiciones en el tercer eje que no se observen en el primer plano factorial? Según lo anterior, ¿vale la pena interpretar el tercer eje?

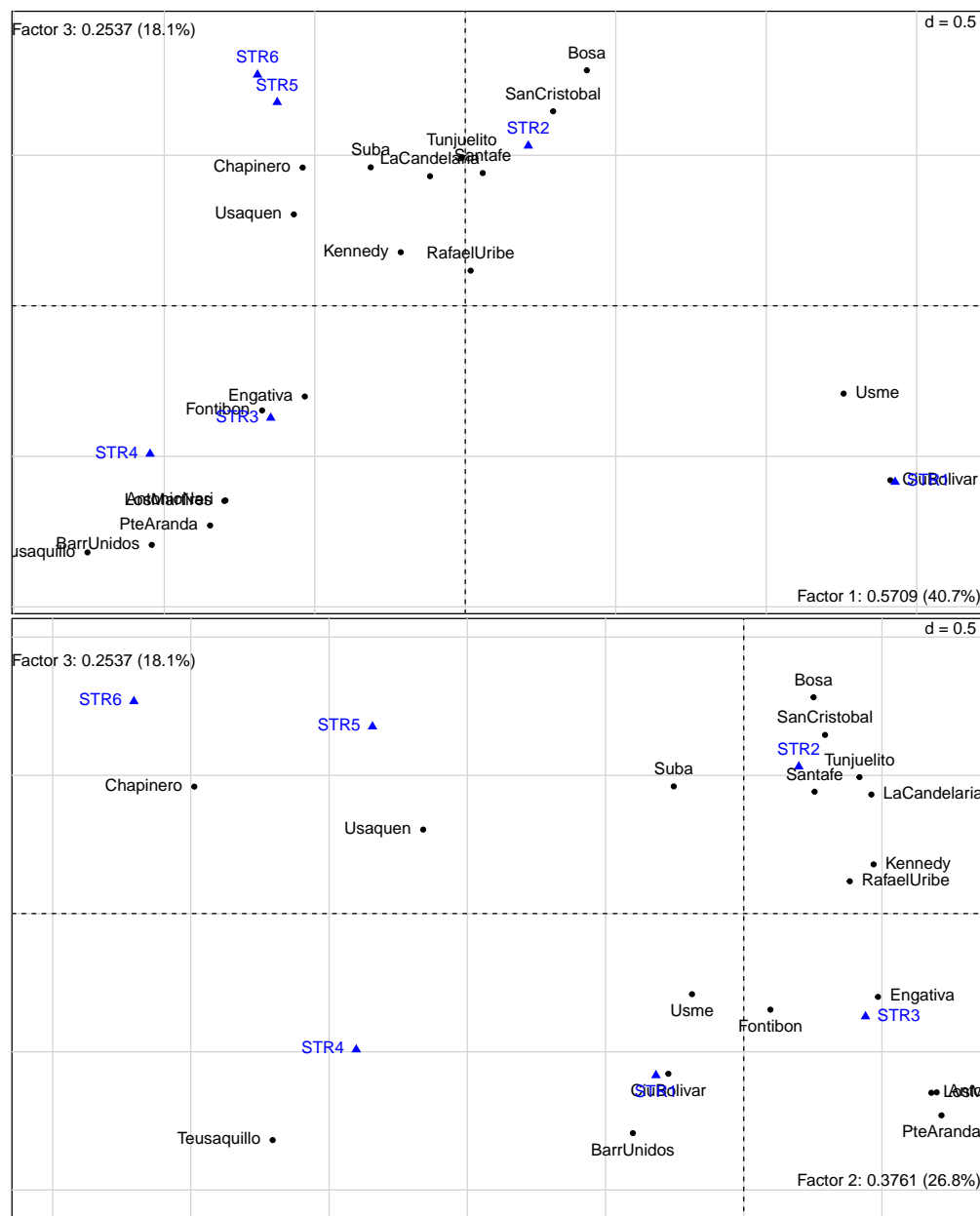


Figura 8: Planos factoriales para los ejes 1 y 3 (arriba), y ejes 2 y 3 (abajo) para la distribución de las manzanas de Bogotá

De acuerdo a la figura 8, en el eje 3 se puede observar que el estrato 4 se contrapone a los estratos 5 y 6, lo cual no era evidente en el primer plano factorial. Sin embargo, este eje no recoge suficiente inercia como para contemplarlo en el análisis. Así mismo, de acuerdo a las tablas de inercia, la localidad Rafael Uribe no estaría bien representada en el eje tres.

15. Agregue a los datos una columna de orden de las localidades según el primer plano factorial. Ordene la TC por esa variable y haga una gráfica que muestre los perfiles de las localidades así ordenadas y el perfil promedio. No incluya la columna sin estrato. Resuma la comparación de los perfiles utilizando esta gráfica y el primer plano factorial.

Observando los perfiles ordenados de acuerdo al primer eje en la Figura 9, aquellos que cuentan con proporciones muy similares de estratificación, en el primer plano factorial se encuentran muy cercanos espacialmente. En particular, las localidades Antonio Nariño, Mártires y Puente Aranda tienen una gran proporción del estrato 3, lo cual es similar a lo observado en el primer plano factorial. Así mismo, en las localidades de Usme y Ciudad Bolívar predomina el estrato 1, lo cual es visible en el primer plano factorial, pues éstas se encuentran en el lado positivo del eje 1 y están contiguas.

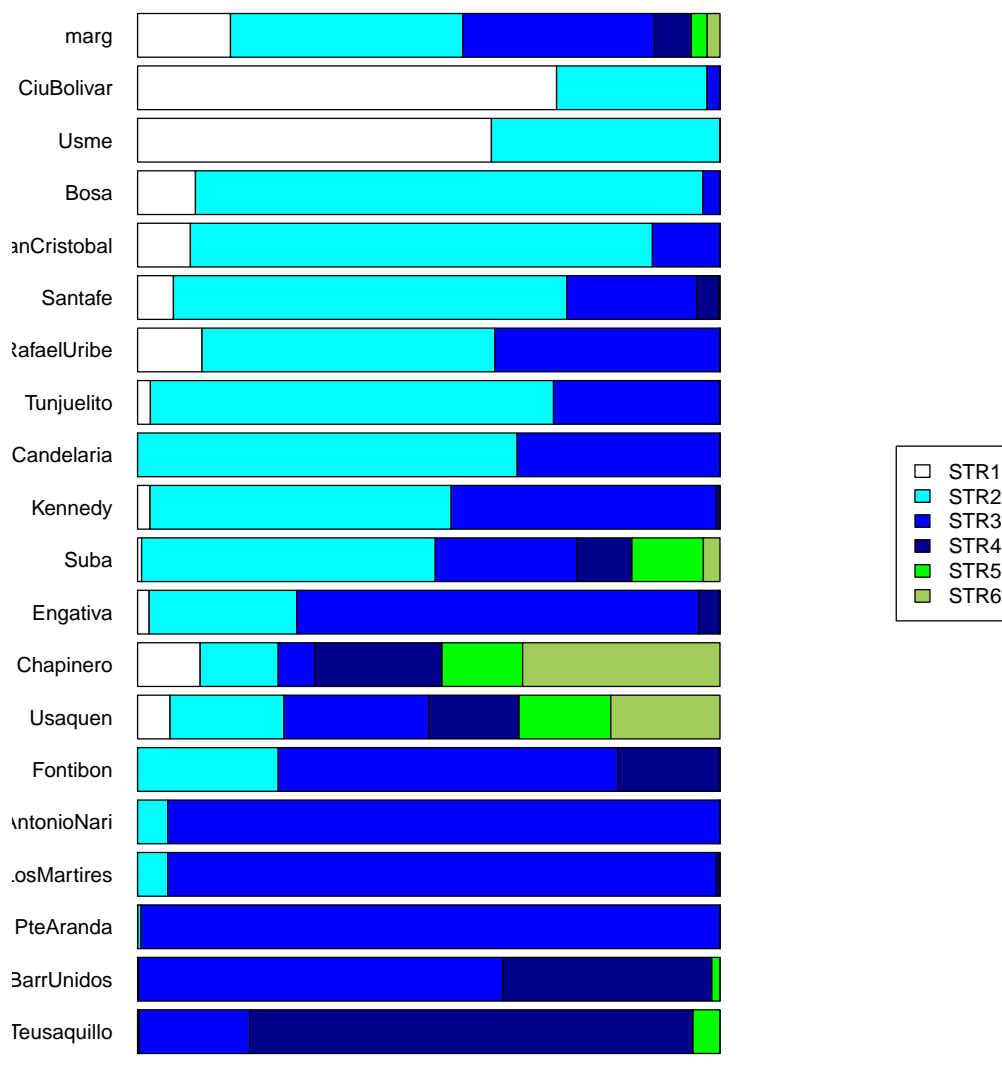


Figura 9: Perfiles de las manzanas de Bogotá ordenadas sobre el primer eje del ACS

16. Utilizando el primer plano factorial proponga una partición de las localidades en cinco clases.

De acuerdo a la distribución de las localidades en el primer plano factorial, la partición de las mismas sería así:

Grupo 1: Usme, Ciudad Bolívar; *Grupo 2:* Kennedy, La Candelaria, Tunjuelito, Rafael Uribe, San Cristóbal, Santafé, Bosa; *Grupo 3:* Puente Aranda, Antonio Nariño, Los Mártires, Engativá; *Grupo 4:* Fontibón, Suba, Barrios Unidos; *Grupo 5:* Usaquén, Teusaquillo, Chapinero.

17. Proponga una partición en cinco clases utilizando la gráfica de perfiles entre localidades.

De acuerdo a la figura 9, los grupos propuestos para las localidades serían los siguientes: *Grupo 1:* Usme, Ciudad Bolívar; *Grupo 2:* Bosa, San Cristóbal, Santafé, Rafael Uribe, Tunjuelito, La Candelaria, Kennedy; *Grupo 3:* Puente Aranda, Antonio Nariño, Los Mártires, Fontibón, Engativá; *Grupo 4:* Barrios Unidos, Teusaquillo; *Grupo 5:* Usaquén, Chapinero, Suba.

18. Proponga una partición “final” de las localidades en cinco clases y haga una gráfica de perfiles incluyendo el perfil marginal. Como otra síntesis del análisis comente la gráfica obtenida.

De acuerdo a la figura 10, las particiones 1, 2 y 3 coinciden con la ubicación de los tres primeros estratos. La partición 5 está asociada a los estratos más altos (4,5,6), mientras que la partición 4 contiene perfiles un poco disímiles pero cercanos espacialmente entre sí.

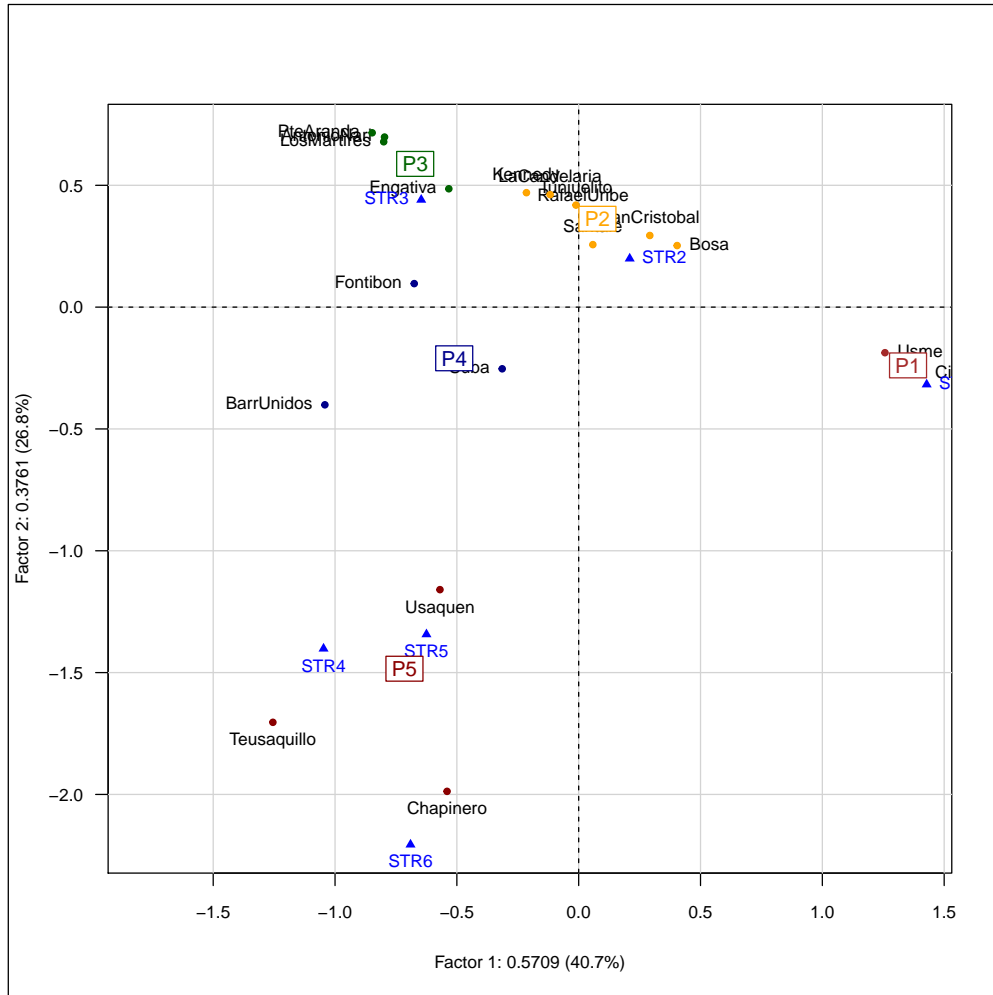


Figura 10: Primer plano factorial del ACS mostrando los centros de gravedad de grupos de localidades (P2 a P5)

19. **Construya una TC más pequeña formando algunos grupos de localidades de perfiles muy parecidos: Usme-Ciudad Bolívar, etc. (las dos filas se colapsan en una sumándolas en la TC). Realice un ACS de esta tabla, compare los resultados de la tabla completa. Comente la propiedad de equivalencia distribucional.**

Al realizar el ACS con la nueva tabla de contingencia del cuadro 4, la cual agrupa perfiles similares, los resultados del ACS original son similares al ACS de las particiones. Observando la figura 11, al agrupar las localidades no cambia significativamente la ubicación de los perfiles en el primer plano factorial. Por tanto, se concluye que en este caso aplica la propiedad de equivalencia distribucional.

Por otro lado, al comparar la inercia acumulada en el ACS original con el ACS de los perfiles agrupados, se encontró que el promedio de la inercia bajó, de 0.28 a 0.26, lo cual implica que hay una pequeña pérdida de variación.

	NoSTR	STR1	STR2	STR3	STR4	STR5	STR6
P1	594	4539	2087	91	0	0	0
Usaquen	327	109	383	487	303	308	367
P2	1205	870	8486	3436	50	1	0
P3	459	1	67	2879	4	0	0
Engativa	334	67	857	2335	122	0	0
Suba	453	27	1982	955	373	482	112
BarrUnidos	98	1	1	675	387	15	0
Fontibon	357	0	296	715	215	0	0
Teusaquillo	148	0	2	156	625	38	0
Chapinero	113	97	121	57	197	125	306

Cuadro 4: Tabla de contingencia según grupos de localidades

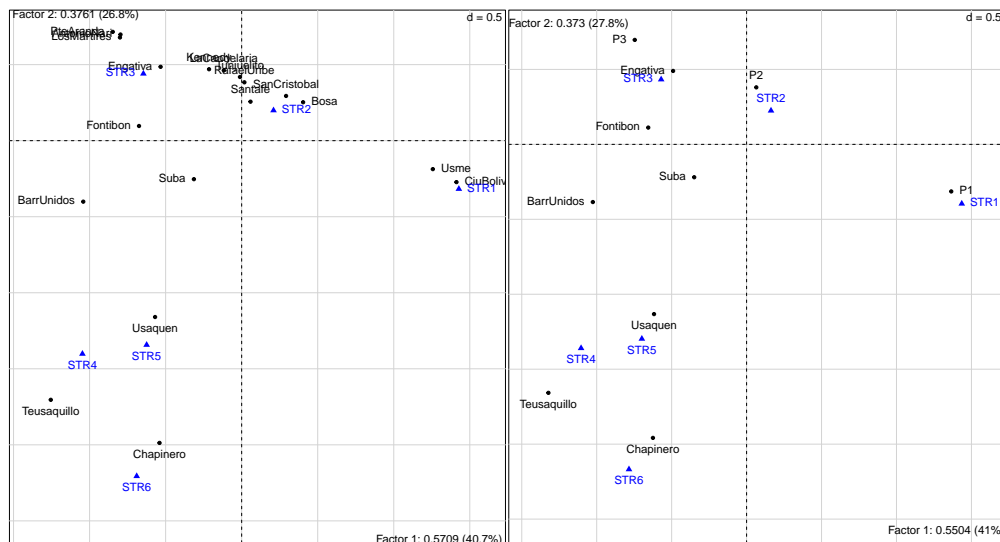


Figura 11: Planos factoriales del ACS original (izquierda) y el ACS de las particiones (derecha)

20. Compruebe “a mano”(utilizando R) las relaciones de transición. Por ejemplo: calcule la coordenada, sobre el eje 1, de Usme a partir de las coordenadas de los 6 estratos. Analice el ejercicio (¿quién atrae a quién y por qué?).

```
> #Relaciones cuasibarcéntricas
> Usmeperf<-tabbc$perR[5,]/100#coordenadas perfil Usme
> eig1<-acs$eig[1] #valor propio 1
> estraxis1<-acs$co[,1] #Coordenada primer eje Usme
> 1/sqrt(eig1)*sum(c(Usmeperf)*estraxis1)
```

```
[1] 1.25782
```

```
> atrac<-Usmeperf*estraxis1
> acs$li[5,]#coordenada Usme primer eje
```

```
Axis1 Axis2 Axis3
Usme 1.257096 -0.1870643 -0.2916999
```

La coordenada del perfil Usme sobre el primer eje es 1.25. Usando las relaciones cuasibaricéntricas, logramos obtener este mismo valor con las coordenadas de los estratos sobre el primer eje. La mayor atracción se presenta con el estrato 1 hacia la localidad de Usme (0.86), puesto que el ponderador (0.60, el perfil de Usme para el estrato 1) es el más alto, y el estrato 1 es el más alejado del origen, hacia el lado positivo del eje 1. Así mismo, gráficamente la relación entre el estrato 1 y Usme es evidente por su cercanía (figura 4). Por otro lado, la dilatación para el perfil de Usme es 1.32, la cual aleja el promedio ponderado (0.95) del centro.

21. Describa la distribución geográfica de los habitantes de Bogotá según su nivel socioeconómico, utilizando el estrato de la manzana donde vive cada uno como indicador de ese nivel.

Una población con capacidad adquisitiva media alta aún puede elegir la zona de vivienda en una localidad específica con dominio de estratos 4 y 5 (e.g. Usaquén y Teusaquillo) contrario a la población de capacidad media baja (estratos 2 y 3) que se encuentra ocupando localidades diversas de estrato. Existe otro conjunto de localidades que están representadas por manzanas de estrato 1 y corresponden a aquellos lugares de fácil acceso para poblaciones con capacidad adquisitiva mínima, es decir cada localidad puede estar caracterizada por la dominancia de un estrato particular. Por otro lado, localidades de niveles socioeconómicos extremos tienen exclusividad de estrato, por ejemplo Chapinero está asociada al estrato 6, mientras que Usme y Ciudad Bolívar están asociadas al estrato 1. Se puede observar que hay una tendencia en la agrupación de localidades a coincidir con una distribución espacial de los estratos en la ciudad, perteneciendo al estrato 1 las localidades ubicadas al sur oriente y centro, al estrato 2 y 3 las localidades ubicadas al sur occidente y 4, 5 y 6 al norte de la ciudad.