

Análisis de la tasa de incidencia del cáncer de piel en Costa Rica mediante métodos de agrupamiento

Aracelly Zárate Cordero¹, Wendy Zárate Cordero¹
aracelly.zarate@ucr.ac.cr, wendy.zarate@ucr.ac.cr

RESUMEN

El presente estudio se enfoca en el análisis de la tasa de incidencia del cáncer de piel en Costa Rica utilizando métodos de agrupamiento. El cáncer de piel es una enfermedad con cada vez más incidencia en todo el mundo, su comprensión y análisis son fundamentales para la implementación de estrategias efectivas de prevención y control. Se emplearon técnicas de agrupamiento con el objetivo de identificar agrupamiento de cantones según la incidencia del cáncer de piel durante los años 2012 y 2015, así como analizar los patrones de incidencia de este tipo de cáncer en el período comprendido entre 2011 a 2016. Los resultados principales proporcionan información valiosa sobre la caracterización de los grupos. Se encontró que los cantones con las tasas de incidencia más altas se concentran principalmente en la Región Central del país. Además, se observó que estos cantones presentan índices de desarrollo humano elevados y una altitud promedio más alta, estos hallazgos resaltan la importancia de considerar factores socioeconómicos y geográficos en el análisis de la incidencia del cáncer de piel en Costa Rica.

PALABRAS CLAVE: enfermedad, cantones, métodos jerárquicos, k-medias

INTRODUCCIÓN

El cáncer es una enfermedad que afecta a millones de personas en todo el mundo. Entre los diferentes tipos de cáncer, el cáncer de piel se ha convertido en uno de los más comunes, y su incidencia ha experimentado un preocupante aumento en diversos países, incluyendo Costa Rica. En este artículo, se aborda el análisis de las tasas de incidencia del cáncer de piel mediante métodos de agrupamiento. Este enfoque permite identificar patrones y características comunes entre los casos de cáncer de piel, lo que puede contribuir a una mejor comprensión de esta enfermedad y a la implementación de estrategias más efectivas para su prevención y tratamiento.

En Costa Rica el cáncer es la segunda causa de muerte, después de las enfermedades cardiovasculares. Según datos del Registro Nacional de Tumores del Ministerio de Salud y del Instituto Nacional de Estadísticas y Censos cada año se presentan en el país 11 500 casos nuevos de cáncer, lo que equivale aproximadamente a 30 personas diagnosticadas con esta enfermedad cada día (Blanco, 2022).

De acuerdo con Echeverría (2008) muchos tumores pueden curarse en etapas tempranas. La implantación de programas de detección temprana y diagnóstico precoz permiten la aplicación de tratamientos que suelen ser más eficaces, menos complejos y más económicos, lo que a su vez puede salvar vidas y reducir la morbilidad en muchas personas. Por su parte, Aparicio y Morera (2009) destacan la importancia de localizar zonas de riesgo, ya que esto constituye una de las

¹Estudiantes de Bachillerato en Estadística, Universidad de Costa Rica

principales estrategias de los sistemas públicos de salud para impulsar programas de intervención focalizados.

El aumento en la incidencia del cáncer puede atribuirse a varios factores, entre los cuales destacan el envejecimiento y crecimiento de la población, los cambios en la prevalencia y distribución de los principales factores de riesgo, así como modificaciones en el acceso a servicios de salud (Piñeros, Vignat, Wiesner y Soerjomataram, 2020). Este último aspecto reviste una importancia fundamental, dado que existe evidencia que relaciona los cambios en la incidencia y mortalidad del cáncer con el nivel de desarrollo socioeconómico de una región.

En un estudio longitudinal llevado a cabo por Leiser et al. (2020) en Estados Unidos, se analizó una cohorte de 260.585 personas de entre 0 y 40 años en 1940. Durante el periodo de seguimiento, que se extendió desde 1940 hasta el diagnóstico de cáncer o la muerte de la persona, se utilizaron estadísticas de exploración espacial. Entre los principales hallazgos del estudio destaca que el área urbana presentó una superposición de múltiples tipos de cáncer, y que dos distritos urbanos en particular tuvieron un mayor riesgo de padecer hasta cinco tipos de cáncer distintos.

En relación con lo anterior, Brenes (2023) señala que, en el caso de Costa Rica, los cantones con mayor nivel socioeconómico tienen tasas más altas de incidencia y mortalidad por cáncer que los cantones con menor nivel socioeconómico. No obstante, en general los cantones de mayor nivel socioeconómico tienen mayor probabilidad de supervivencia después del diagnóstico del cáncer. Del mismo modo, se han identificado patrones de alta incidencia en el Pacífico Central, especialmente en cáncer de piel, colon y próstata y en la provincia de Cartago, donde se registran tasas elevadas de cáncer de estómago, órganos digestivos y genitourinarios (CCP, 2023).

Por otro lado, se han identificado patrones de baja incidencia en la Zona Sur y Zona Norte del país, especialmente en cáncer de próstata y mama, así como en cáncer en órganos digestivos y genitourinarios en mujeres y de colon y tiroides en ambos sexos. Con el análisis de variabilidad se encontró que existen diferencias de más de tres veces en la incidencia de cáncer entre áreas de salud (CCP, 2023).

Estos hallazgos están respaldados por el estudio de Murillo y González (2020), quienes utilizaron el índice de Thiel para cuantificar la desigualdad en la incidencia y mortalidad del cáncer de próstata en los distintos cantones del país. Para el año 2015, se observó un incremento en los nuevos casos, pero no se distribuyó de manera uniforme en todo el país, sino que se concentró en algunos cantones en particular. Sin embargo, debido a la metodología empleada, no se pudieron identificar específicamente cuáles fueron esos cantones.

Particularmente, el cáncer de piel presenta diversos factores de riesgo identificados en su incidencia, tales como la edad, el sexo, el fenotipo de la piel, la exposición solar, la ocupación, entre otros (Rodríguez, Granda, Hechavarría y Azze, 2003). Según Sánchez (2006), la exposición a la radiación solar es el factor de riesgo más relevante en el desarrollo del cáncer cutáneo. Además, Pascuale et al. (2013) señala que la radiación solar aumenta un 15% por cada 1000 metros de altitud. Esta información indica que las regiones con mayor altitud presentan un mayor riesgo de padecer esta enfermedad.

Con base en lo expuesto anteriormente, el objetivo principal de esta investigación es analizar la incidencia del cáncer de piel en 81 cantones de Costa Rica durante los años 2012 y 2015, para identificar agrupamientos. Asimismo, se pretende analizar los patrones de incidencia del cáncer de piel en el período comprendido entre 2012 y 2016 en estos mismos cantones. Se busca identificar patrones y agrupamientos de áreas geográficas que presenten características similares en términos de la incidencia de este tipo de cáncer.

METODOLOGÍA

Los datos utilizados en este estudio fueron proporcionados por el Ministerio de Salud Pública y se presentaron en formato de tabla estadística. Para su análisis, fue necesario transformar estos datos en una base de datos estructurada. La base de datos utilizada contiene información detallada sobre la incidencia de los 10 tipos de cáncer más comunes en Costa Rica, desglosados por cantón y que fueron diagnosticados durante el periodo comprendido entre 2011 y 2016.

La unidad de estudio es el cantón. A partir de los casos incidentes se estima la tasa de incidencia de cáncer de piel tanto para hombres como para mujeres como se muestra en la ecuación (1):

$$\frac{\text{Número de nuevos casos de cáncer}}{\text{Población en riesgo}} * 100\ 000 \quad (1)$$

El número de casos nuevos de cáncer de piel por cantón se obtiene a partir del Ministerio de Salud Pública, mientras que la población en riesgo se define como la población total del cantón y se obtiene a partir de las proyecciones demográficas realizadas por el Centro Centroamericano de Población (CCP). En el contexto de esta investigación, se cuenta con datos correspondientes a 81 de los 84 cantones actualmente existentes en el país. Es relevante destacar que la información disponible abarca únicamente los 81 cantones que estuvieron establecidos durante el período de estudio, excluyendo aquellos que fueron creados recientemente.

Para llevar a cabo el agrupamiento de los datos, se utilizaron como variables la tasa de incidencia de cáncer de piel en hombres y mujeres, desglosada por cantón. Además, se emplearon tres variables adicionales para caracterizar los grupos identificados: la altitud promedio del cantón, el índice de desarrollo humano cantonal y la población adulta mayor del cantón (personas mayores de 65 años). Estas variables complementarias se seleccionaron con el propósito de obtener una visión más completa y detallada de los patrones y características asociadas a la incidencia del cáncer de piel en cada cantón. Se emplearon métodos jerárquicos con el fin de identificar patrones de similitud y diferencia en la incidencia de cáncer de piel.

En primer lugar, se aplicaron métodos jerárquicos para examinar los agrupamientos de cantones durante los años 2012 y 2015. En este caso, se utilizó la distancia euclídea como medida de similitud entre los cantones. Esta medida cuantifica la diferencia absoluta entre las tasas de incidencia de los cantones, permitiendo identificar aquellos que presentan similitudes en términos de la incidencia de cáncer de piel. Adicionalmente, se eligió el método de “vecino más lejano” como medida de distancia entre grupos.

Por otro lado, se examinó la tendencia de las tasas de incidencia en cada cantón durante el periodo de 2011 a 2016. Para este análisis, se empleó la distancia de correlación, que es una medida de similitud que evalúa la forma o patrón de las variables. Esta medida permite identificar aquellos cantones que presentan patrones de incidencia similares a lo largo de los años, es decir, aquellos que muestran tendencias de aumento o disminución en la incidencia de cáncer de piel de manera similar. Además, se utilizó la medida de salto promedio como distancia entre grupos en el proceso de agrupamiento.

El uso de la distancia euclídea y la distancia de correlación en distintos análisis se fundamenta en la naturaleza de los datos, los objetivos específicos de cada análisis y considerando el coeficiente de correlación copenhético presentado en el anexo (Tabla 3 y Tabla 4). La distancia euclídea se emplea para identificar agrupamientos en un punto temporal específico, mientras que la distancia de correlación se utiliza para analizar la tendencia temporal de las tasas de incidencia.

Utilizando las medidas de distancia mencionadas, se procedió a realizar la representación gráfica del agrupamiento jerárquico de los datos mediante un dendrograma. Este gráfico permite visualizar las relaciones de similitud entre los cantones en función de las variables consideradas. Una vez obtenido el dendrograma, se procede a definir los clústeres mediante el corte de las ramas del dendrograma. Para determinar la cantidad adecuada de clústeres se utilizan gráficos de silueta promedio y el criterio de suma de cuadrado dentro de grupo (SCDG) el cual mide la dispersión de los datos dentro de cada grupo o clúster, el objetivo es encontrar el número de clústeres que minimice la SCDG, lo que indica una mayor coherencia interna y una mejor separación entre los grupos.

Finalmente, se aplicó el algoritmo de agrupamiento de k-medias. A partir de los clústeres obtenidos mediante el dendrograma, se calculó el centroide de cada clúster. Luego, utilizando estos centroides como puntos de partida, se asignó cada cantón al clúster cuyo centroide fuera el más cercano. Se repitieron iterativamente los pasos de cálculo de distancias y actualización de centroides hasta que no se observaron cambios en la asignación de los cantones a los clústeres.

El análisis se realizó mediante el lenguaje de programación R (R Core Team, 2021) utilizando la versión 4.1.1 con las librerías biotools (da Silva, 2021), cluster (Maechler et al., 2022), dendextend (Galili et al., 2023), factoextra (Kassambara y Mundt, 2020), y ggplot2 (Wickham, 2016).

RESULTADOS

Para comprender el comportamiento inicial de los datos, resulta fundamental proporcionar contexto mediante algunas características de las variables utilizadas. En este sentido, en la Tabla 2 (ver anexo), se muestra el promedio y la varianza de dichas variables por año y sexo. Al analizar estas medidas, se observa que las varianzas de las variables difieren considerablemente, lo que indica la necesidad de estandarizar las variables antes de examinar la tendencia de las tasas de incidencia.

A su vez, para identificar valores atípicos se emplea la inspección gráfica que se muestra en la Figura 10 (ver anexo), se decide eliminar únicamente aquellas observaciones que se alejan

considerablemente del límite establecido (línea roja). Por otro lado, se optó por mantener el resto de las observaciones, aunque estén por encima del límite, debido a que estas observaciones podrían contener información valiosa sobre la tasa de incidencia.

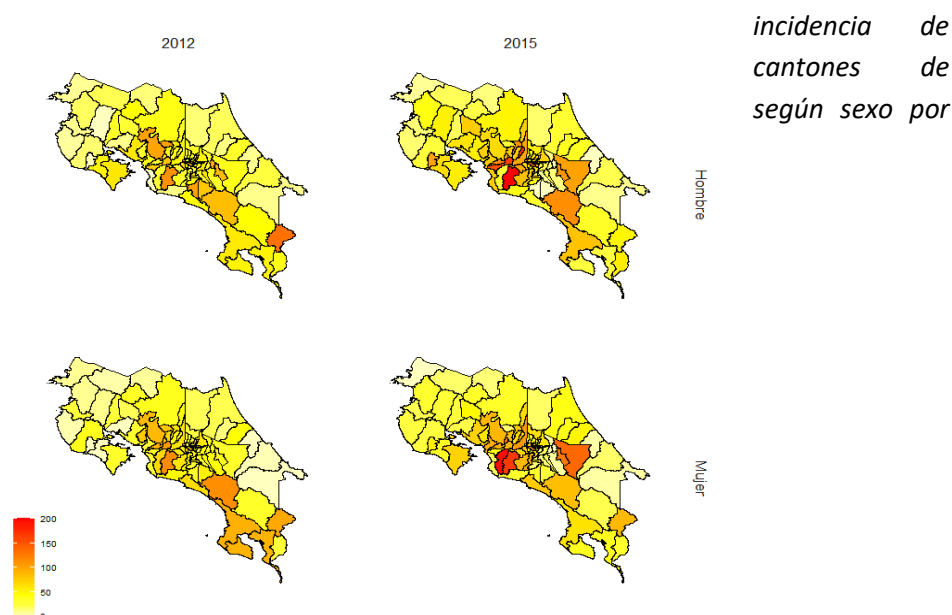
En relación con lo anterior, en la Figura 10 (ver anexos) se presentan los valores de leverage para las tasas de incidencia de los años 2012 y 2015. Para el 2012 no se eliminan observaciones, en el año 2015 se elimina la información procedente del cantón de Turrubares. Por otro lado, en la Figura 11 (ver anexos) se muestra el resultado de los valores leverage para la tasa de incidencia masculina y femenina en el período de tiempo de 2011 a 2016. Para realizar el análisis de tendencia se eliminan la observación de Abangares y Puriscal en las tasas de incidencia masculina y femenina respectivamente.

En la Figura 1, se presentan las tasas de incidencia desglosadas por año, sexo y por 81 cantones del país. Se utiliza una escala de colores donde el tono más intenso representa la tasa más alta de incidencia. Al analizar el gráfico de manera general, se observa que los cantones presentan una mayor intensidad en el año 2015. En comparación con años anteriores, los cantones que anteriormente tenían tonalidades más tenues han experimentado un cambio hacia tonalidades más intensas, este cambio en la intensidad del color indica un incremento en las tasas de incidencia a lo largo del tiempo.

Estos hallazgos sugieren que las tasas de incidencia han ido en aumento, de manera específica se observa que para el 2012 las mayores tasas de incidencia en hombres se presentaron en Coto Brus, Puriscal, Zarcero y Tarrazú. Mientras que, en el caso de las mujeres, los cantones con mayores tasas fueron Puriscal, Pérez Zeledón y Atenas. En el año 2015 las tasas más altas se presentaron en Puriscal, Poas, Atenas, Orotina y San José, en el caso de los hombres, los cantones de Turrubares, Puriscal, Turrialba y Palmares tuvieron las tasas de incidencia más alta para las mujeres.

Figura 1

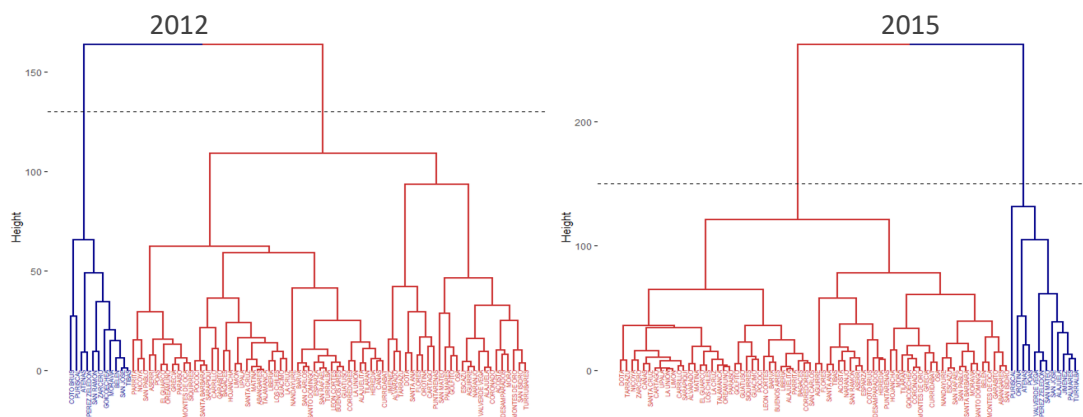
*Tasas de
cáncer de piel en
Costa Rica
año*



A su vez, en la Figura 2, se muestra el resultado del agrupamiento inicial de los cantones utilizando métodos jerárquicos con la distancia euclídea y “vecino más lejano”. Mediante este análisis, se han formado dos grupos de cantones. La formación de estos grupos se ha llevado a cabo teniendo en consideración el criterio de SCDG y gráficos de silueta promedio presentados en los anexos (Figura 8). En ambos años ocurre que el grupo 1 está conformado por la gran mayoría de cantones, 70 y 68 cantones para los años de 2012 y 2015 respectivamente, mientras que, el grupo 2 está conformado solo por 11 cantones en el 2012 y 12 cantones en el año 2015.

Figura 2

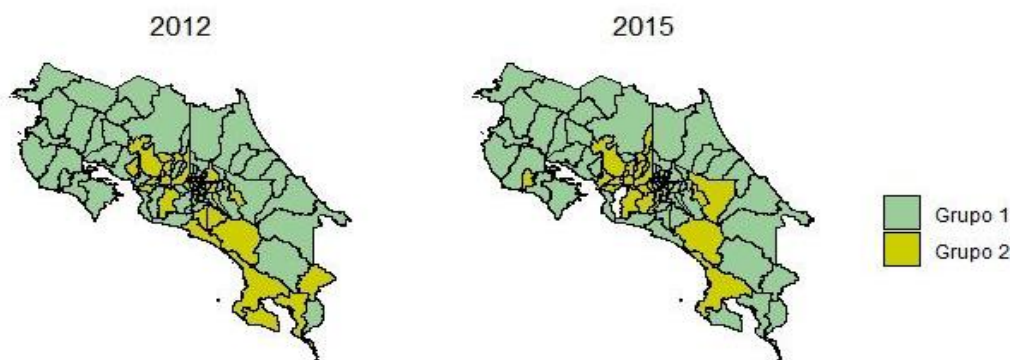
Clasificación Jerárquica de cantones con la distancia euclídea y vecino más lejano según año.



Después de aplicar el algoritmo de K-medias, se obtienen los agrupamientos que se muestran en la Figura 3. En el año 2012, el grupo 1 ahora está compuesto por 51 cantones, mientras que el grupo 2 ha aumentado de 11 a 30 cantones. De manera similar, en el año 2015, el primer grupo está conformado por 56 cantones y el grupo 2 ha pasado a tener 24 cantones. En general, se puede observar que los agrupamientos de los cantones han experimentado cambios de un año a otro. Sin embargo, se puede observar que los cantones de la región central del país tienden a mantenerse principalmente en el grupo 2, independientemente del año.

Figura 3

Agrupamiento de los cantones después de aplicar el algoritmo de k-medias, según año.



Se calculó la suma de cuadrados dentro de grupo para cada año y para ambos métodos de agrupación (agrupamiento jerárquico y k-medias). En ambos años, el método de k-medias mostró un valor más pequeño. Esto sugiere que el agrupamiento obtenido con k-medias es el más adecuado, y es este agrupamiento el que se utiliza para realizar la caracterización de los grupos.

A través de la caracterización de los grupos, se puede observar en la Figura 4 que tanto para el año 2012 como para el año 2015, los cantones con las tasas más altas de incidencia para hombres y mujeres se encuentran en el grupo 2. Al examinar el Índice de Desarrollo Humano (IDH), se puede notar que los cantones en el grupo 2 presentan una variabilidad similar en comparación con los del grupo 1. Sin embargo, se destaca que la mediana del grupo 2 es ligeramente más alta que la mediana del grupo 1, lo cual indica que los cantones en el grupo 2 tienen índices de desarrollo humano un poco más altos.

De forma similar para ambos años, al inspeccionar la altitud promedio se puede notar que el grupo 1 presenta una mayor variabilidad, es notable que los cantones pertenecientes al grupo 2 son aquellos que presentan las mayores altitudes promedio en comparación con el grupo 1.

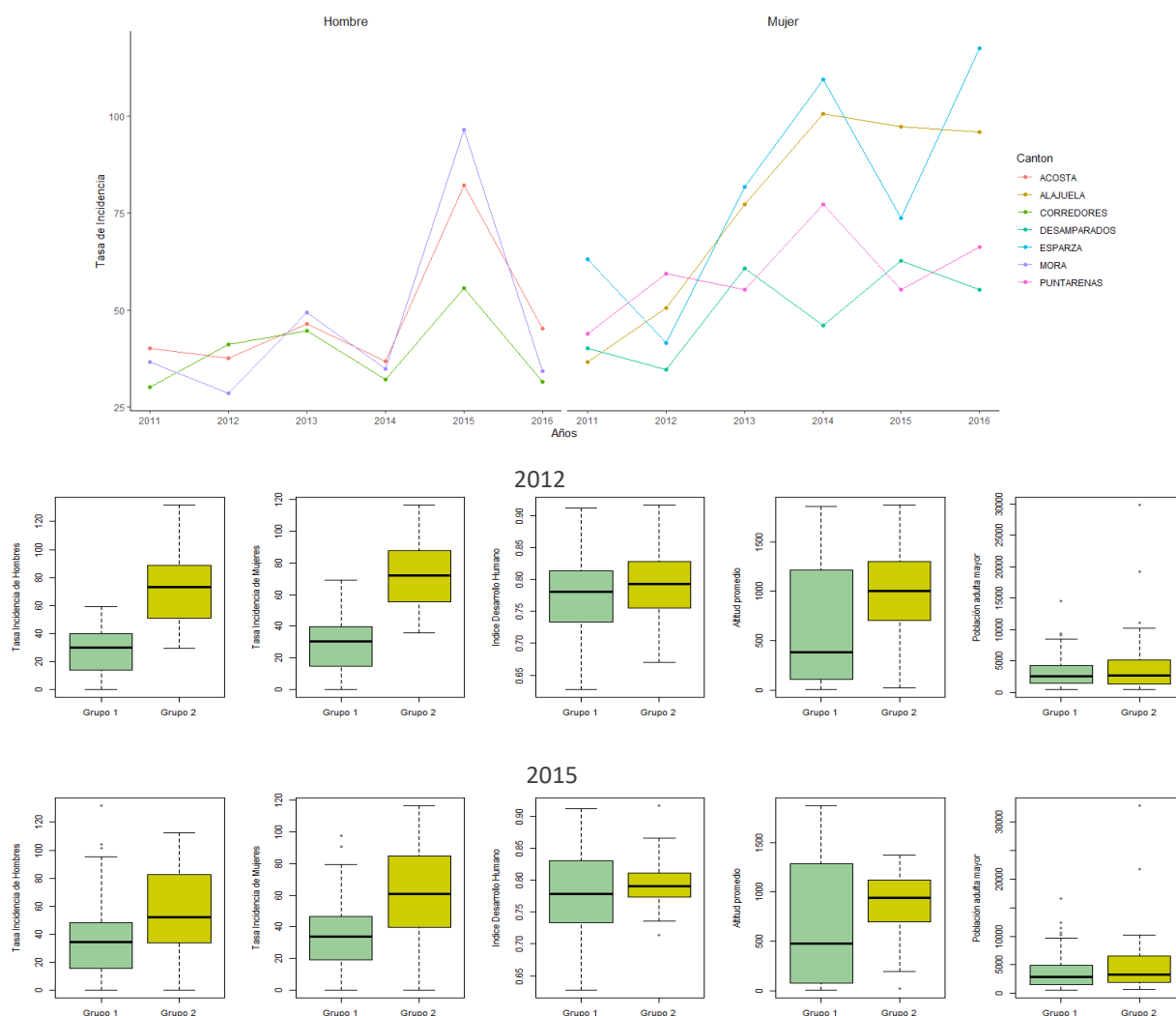
En cuanto a la población adulta mayor (mayores de 65 años), ambos grupos muestran una variabilidad similar y las medianas son comparables, esto indica que los valores centrales de ambos grupos son similares. Sin embargo, es importante destacar que en el grupo 2 se observan valores extremos, lo que implica que existen cantones con una gran cantidad de población adulta mayor y altas tasas de incidencia. Este hallazgo sugiere que la presencia de una gran población adulta mayor puede estar relacionada con mayores tasas de incidencia en ciertos cantones del grupo 2.

Figura 4

Caracterización de grupos de cantones según Tasa de Incidencia por Género, IDH, Altitud Promedio y Población Adulta Mayor.

A continuación, se presentan los resultados del análisis de tendencia de las tasas de incidencia durante los años 2011 a 2016. En la Figura 5 se muestra la tendencia de los cantones que presentaron una alta correlación y comparten un patrón similar en su tendencia. En el caso de la tasa de incidencia en hombres, los cantones de Acosta, Corredores y Mora exhiben un patrón similar a lo largo de los años. Asimismo, en la tasa de incidencia en mujeres se observan similitudes entre los cantones de Puntarenas y Esparza, así como entre los cantones de Desamparados y Alajuela.

Figura 5
Tendencia de las tasas de incidencia por cáncer de piel, según sexo durante el periodo 2011-2016.



De igual forma, se considera el criterio SCDG y gráficos silueta promedio (Figura 9, en anexos), se definen 2 conglomerados y se utilizan métodos jerárquicos con la medida de correlación y el salto promedio como se presentan en la Figura 12 (en anexos). Los resultados después de aplicar el algoritmo de k medias se presentan en la Figura 6, con ayuda de la Tabla 1, se observa que la tendencia de la tasa de hombres, el grupo 1 (donde se registran las mayores tasas) está compuesto principalmente por los cantones de las regiones Central, Brunca y Pacífico Central. Por otro lado, en el caso de las mujeres, los cantones con las mayores tasas y con una tendencia similar se encuentran principalmente en el grupo 2, conformado por cantones de las regiones Central y Pacífico Central.

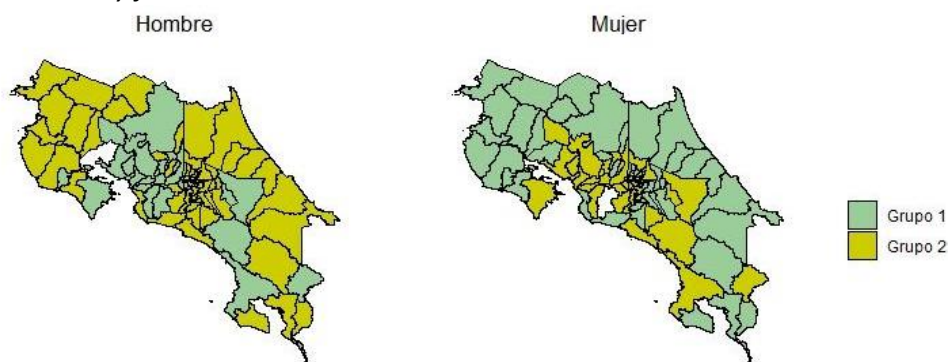
Tabla 1

Agrupación de cantones por tasa de incidencia masculina y femenina según la región geográfica

Región	Tasa Masculina		Tasa Femenina	
	Grupo 1	Grupo 2	Grupo 1	Grupo 2
Brunca	3	3	3	3
Central	22	23	16	28
Chorotega	2	8	10	1
Huetar Caribe	0	6	6	0
Huetar Norte	1	4	5	0
Pacífico Central	5	3	2	6

Figura 6

Agrupamiento de los cantones después de aplicar el algoritmo de k-medias, según tasa de incidencia masculina y femenina



Se realiza el cálculo de la suma de cuadrados dentro de grupo, para cada sexo y con ambos métodos de agrupamiento. Los resultados indican claramente que el método de k-medias es el más adecuado, ya que muestra valores más bajos en la SCDG en comparación con el método jerárquico, de esta forma la caracterización se realiza con el agrupamiento obtenido con k-medias.

Al realizar la caracterización de los grupos con las variables índice de desarrollo humano, altitud promedio del cantón y la población de adultos mayores, se observa en la Figura 7 que en el caso de la tasa de incidencia masculina los cantones del grupo 1 son los que presentan mayor tasa de incidencia, con respecto al índice de desarrollo humano este también es mayor en el grupo 1, si se observa la variable altitud la mediana es mucho mayor en comparación con el grupo 2, por último

la variable de población adulta mayor son muy similares pero es posible observar que en el grupo 1 existe una mayor variabilidad.

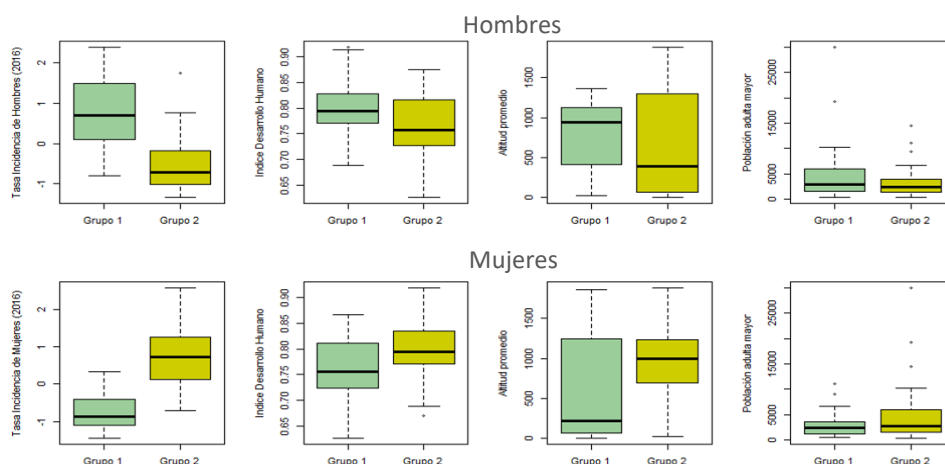
Ahora con respecto a las mujeres los cantones con mayor incidencia se agruparon en el grupo 2, y nuevamente se puede observar que el índice de desarrollo humano es mayor en el grupo 2, al igual que la variable altitud y por último con respecto a la población adulta mayor la mayor variabilidad se presenta en el grupo de mayor incidencia de la enfermedad, aunque las medianas en ambos grupos son muy similares.

Figura 7
Caracterización de grupos de cantones según Tasa de Incidencia por Género, IDH, Altitud Promedio y Población Adulta Mayor

CONCLUSIÓN

En conclusión, el análisis de agrupamiento de los cantones no solo ayuda a identificar aquellos con características similares, sino que también proporciona información valiosa sobre las características distintivas de cada grupo. Este estudio ha revelado la formación de dos clústeres entre los 81 cantones del país, y además se ha demostrado que existen diferencias entre estos grupos según las variables utilizadas para su caracterización, específicamente la zona geográfica, la altitud promedio y el índice de desarrollo humano.

Al analizar la tendencia de los cantones, se confirma que son principalmente los cantones de la región Central los que mantienen constantemente en el grupo con las tasas más altas de incidencia de cáncer de piel a lo largo de los años. Este hecho se relaciona directamente con la información proporcionada por Mayorga (2022), quien destaca que en el Valle Central se alcanzan niveles de "exposición muy alta" a los rayos ultravioleta según el Índice Ultravioleta.



En relación con lo mencionado anteriormente, se puede constatar que los cantones con una altitud promedio más elevada también exhiben las tasas de incidencia más altas. Esto concuerda con lo mencionado por Rodríguez et al. (2003) y Pascuale et al. (2013), quienes mencionan que, a mayor altitud, mayor será la exposición solar, y esto es un factor clave en la aparición de cáncer de piel según lo señalado por Sánchez (2006).

En cuanto al Índice de desarrollo humano, se observó que es más alto en los cantones con mayor tasa de incidencia. Esto concuerda en cierta medida con los resultados obtenidos por Brenes (2023), quien encontró que un mayor nivel socioeconómico se relacionaba con los cantones con las tasas más altas de incidencia de distintos tipos de cáncer. Finalmente, en contradicción con lo expuesto por Piñeros et al. (2020), no se encontraron diferencias entre el agrupamiento de los cantones al caracterizarlos según la población adulta mayor que habitaba en cada cantón.

En conclusión, la detección de patrones de incidencia de cáncer de piel es de vital importancia para comprender la distribución y los factores de riesgo asociados a esta enfermedad. Sin embargo, es importante tener en cuenta las limitaciones de utilizar solo un tipo de cáncer en los estudios epidemiológicos.

Al centrarse únicamente en un tipo de cáncer de piel, se puede perder de vista la complejidad y la interacción de diferentes factores de riesgo que pueden estar involucrados en el desarrollo de dicha enfermedad. Además, es importante reconocer que otros factores ambientales y genéticos también pueden influir en la incidencia del cáncer de piel y deben ser tomados en cuenta en futuras investigaciones.

Es importante considerar que, aunque la detección de patrones de incidencia de un tipo de cáncer de piel es valiosa, es esencial considerar las limitaciones y adoptar un enfoque más amplio que abarque diferentes tipos de cáncer de piel y otros factores de riesgo relevantes para una comprensión más completa de esta enfermedad y para la implementación de medidas de prevención y control adecuadas.

BIBLIOGRAFÍA

- Aparicio, A., & Morera, M. (2009). Patrones geográficos de incidencia y mortalidad evitable por cáncer de mama en Costa Rica. *Revista Costarricense Salud Pública*, 18(01), 43-47. <https://www.scielo.sa.cr/pdf/rcsp/v18n1/a08v18n1.pdf>
- Blanco, P. (28 de septiembre de 2022). Una nueva esperanza en la lucha contra el cáncer. *Universidad de Costa Rica Noticias: Ciencia + Tecnología*. [https://www.ucr.ac.cr/noticias/2022/09/28/una-nueva-esperanza-en-la-lucha-contra-el-cancer.html#:~:text=El%20c%C3%A1ncer%20es%20la%20segunda,Estad%C3%ADsticas%20y%20Censos%20\(INEC\)](https://www.ucr.ac.cr/noticias/2022/09/28/una-nueva-esperanza-en-la-lucha-contra-el-cancer.html#:~:text=El%20c%C3%A1ncer%20es%20la%20segunda,Estad%C3%ADsticas%20y%20Censos%20(INEC)).
- Brenes, G. (2023). Inequidades geográficas en cáncer: ¿Un patrón diferente al de otras enfermedades? *Estadísticas Publicas*. <https://ccp.ucr.ac.cr/rec-info/blog-estadisticas-publicas/entrada-7>
- Centro Centroamericano de Población (CCP). (2023, 02 de febrero). *Introducción: Incidencia de Cáncer*. Atlas de Salud. https://ccp.ucr.ac.cr/farmacoeconomia/atlas/intro_incidencia.htm
- Da Silva, A., Malafaia, G. y Menezes, I. (2017). biotools: an R function to predict spatial gene diversity via an individual-based approach. *Genetics and Molecular Research*, 16.
- Echeverria, J. (2008). Importancia de la detección temprana de cáncer. *Revista Médica Herediana*, 19(04), 135-137. <http://www.scielo.org.pe/pdf/rmh/v19n4/v19n4e1.pdf>
- Galili, T. (2015). dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btv428>
- Kassambara, A. y Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. <https://CRAN.R-project.org/package=factoextra>
- Leiser, C., Taddie, M., Hemmert, R., Richards, R., VanDerslice, J., Henry, K., Ambrose, J., O'Neil, B., Smith, K., & Hanson, A. (2020). Spatial Clusters of Cancer Incidence: Analyzing 1940 Census Data Linked to 1966-2017 Cancer Records. *Cancer Causes Control*, 31(7), 609-615. <https://doi.org/10.1007/s10552-020-01302-3>
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. y Hornik, K. (2021). cluster: Cluster Analysis Basics and Extensions.
- Mayorga, A. (01 de enero de 2022). Rayos ultravioletas alcanzan 'exposición alta' en Costa Rica. *La Nación*. <https://www.nacion.com/el-pais/salud/rayos-ultravioleta-alcanzan-exposicion-alta-en/GHNFHMC4D5HFLFNZVPVZDXS3XW4/story/>
- Murillo, A., & González, M. (2020). *Análisis espacial y temporal de la incidencia y mortalidad por cáncer de próstata y su relación con el acceso geográfico a los servicios de salud por cantón en Costa Rica, 1985 a 2016*. [Tesis de licenciatura, Universidad de Costa Rica]. Repositorio Institucional de la Universidad de Costa Rica. <https://www.kerwa.ucr.ac.cr/handle/10669/83153>

- Pascuale, A., Cordero, A., Núñez, D., Fernandez, C., & Signorelli, I. (2013). Determinación de la influencia de la altitud en el Índice de Protección Natural (IPN) de la piel y el Factor de Protección Solar requerido (SPF). *Revista del Instituto Nacional de Higiene Rafael Rangel*, 44(02), 12-20. <https://pesquisa.bvsalud.org/portal/resource/pt/lil-746322>
- Piñeros, M., Vignat, J., Wiesner, C., & Soerjomataram, I. (2020). La carga del cáncer a nivel mundial: patrones para la toma de decisiones. *Medicina*, 42(4), 640-655. <https://doi.org/10.56050/01205498.1565>.
- R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Rodríguez, R., Granda, A., Hechavarría, J., & Azze, M. (2003). Percepción del riesgo de cáncer de piel por exposición solar. A propósito de una experiencia en Cuba. *Salud de los Trabajadores*, 11(01), 75-77.
- Sánchez, F. (2006). Consideraciones sobre la capa de ozono y su relación con el cáncer de piel. *Revista Médica de Chile*, 134(09), 1185-1190. <http://dx.doi.org/10.4067/S0034-98872006000900015>
- Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag*.

ANEXOS

Tabla 2

Media y varianza de las tasas de incidencia por sexo según año.

	Tasa Masculina					Tasa Femenina				
	2012	2013	2014	2015	2016	2012	2013	2014	2015	2016
Media	44.20	48.79	49.76	58.08	62.93	44.38	50.23	49.08	52.18	58.28
Varianz	849.0	831.3	1150.4	1699.2	2227.5	834.3	1129.3	1678.0	1480.6	1624.5
a	1	2	4	1	7	7	3	5	6	7

Tabla 3

Coefficiente de correlación copenético por año y medida de distancia entre grupos según la medida de distancia entre individuos.

Distancia entre individuos	Vecino más cercano		Vecino más lejano		Salto promedio		Ward	
	2012	2015	2012	2015	2012	2015	2012	2015
Euclídea	0.66	0.73	0.73	0.77	0.75	0.80	0.60	0.54
Mahalanobis	0.52	0.84	0.64	0.84	0.65	0.89	0.56	0.41
Manhattan	0.63	0.69	0.62	0.71	0.73	0.81	0.60	0.50

Tabla 4

Coefficiente de correlación copenético de la distancia de correlación por tasas de incidencia masculina y femenina, según la medida de distancia entre individuos.

	Hombres	Mujeres
Vecino más cercano	0.22	0.50
Vecino más lejano	0.54	0.56
Salto promedio	0.64	0.67
Ward	0.54	0.62

Figura 8

Silueta promedio para el agrupamiento de cantones según año.

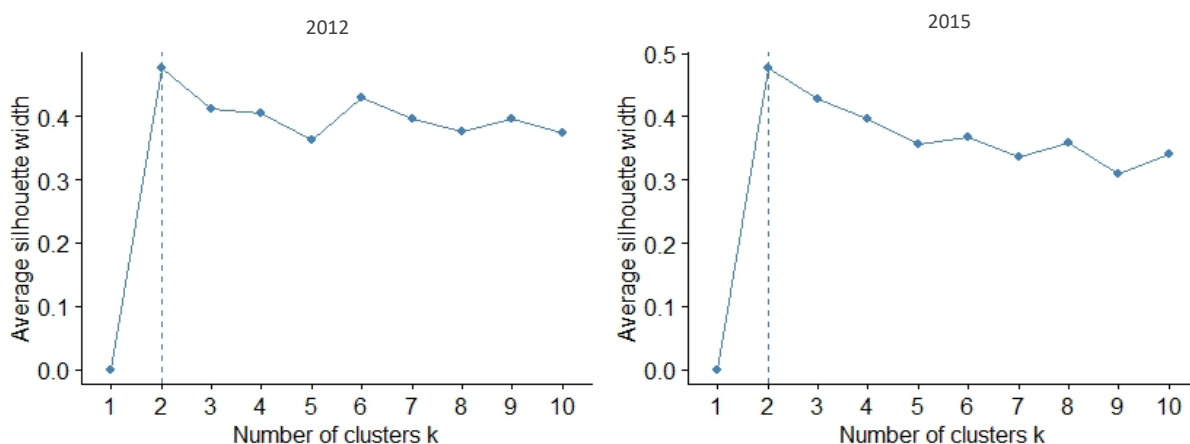


Figura 9

Silueta promedio para el agrupamiento de cantones según la tasa de incidencia masculina y femenina.

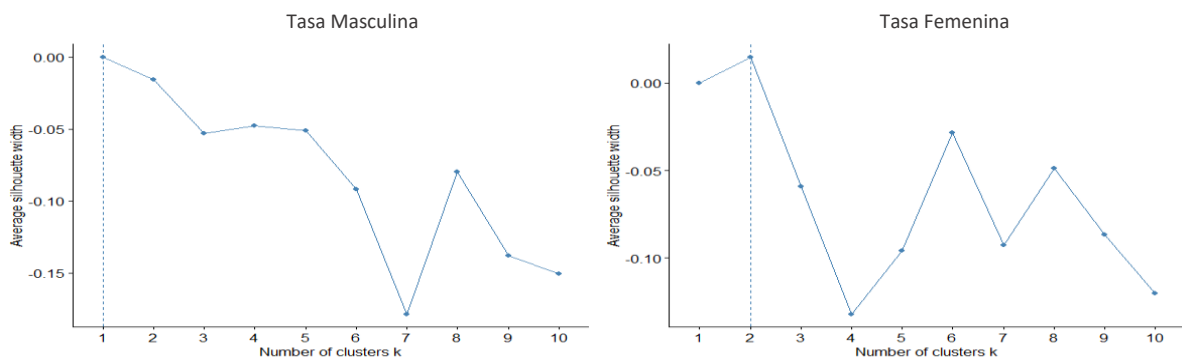


Figura 10

Análisis de valores leverage de las tasas de incidencia de cáncer de piel, según año.

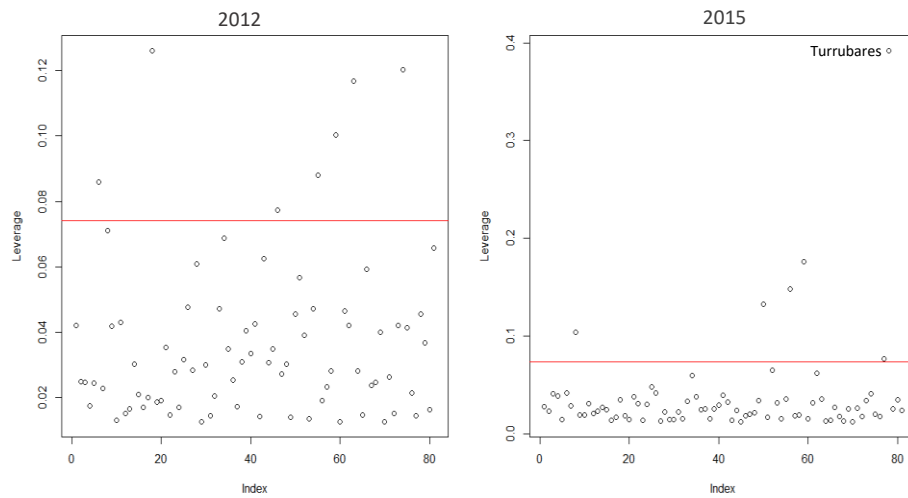


Figura 11

Análisis de valores leverage de las tasas de incidencia masculina y femenina.

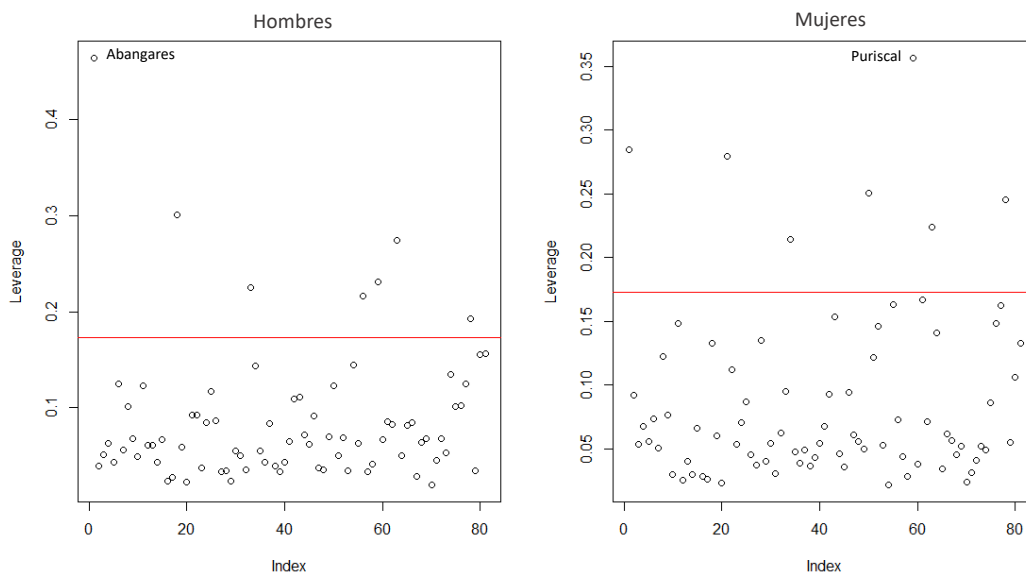


Figura 12

Clasificación Jerárquica de cantones con la distancia de correlación y salto promedio según tasa de incidencia masculina y femenina.

