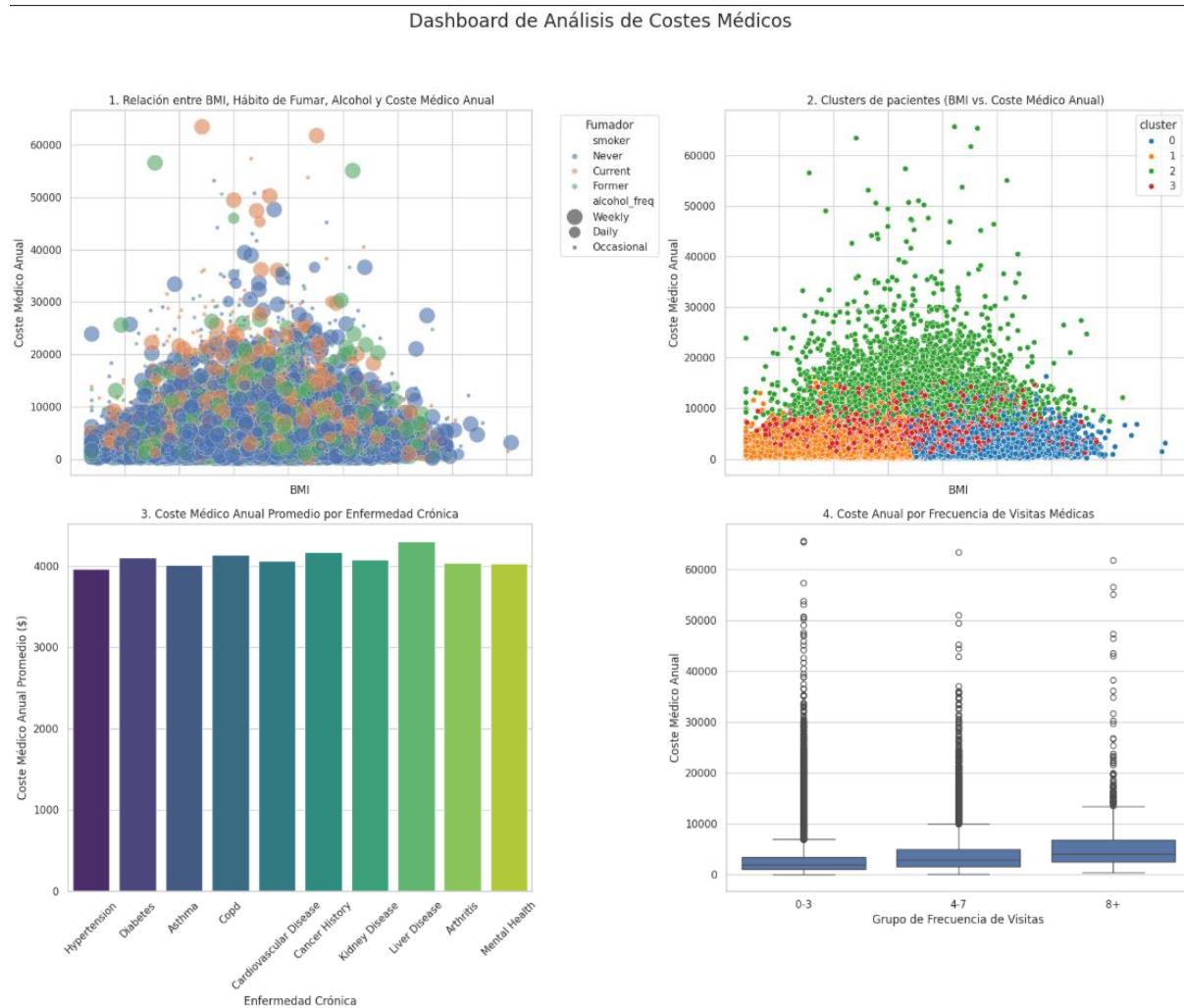


**Nombre y Apellidos: Gabriel Llaneza**

**Github con notebook:** <https://github.com/Gabriellcai/AI-Genre-Detector-LSTM/tree/main>

## 1. Resumen Ejecutivo.

En este dashboard incluyo las gráficas y conclusiones claves que he sacado del análisis detallado de modelo de negocio de la aseguradora.



- Hábitos:** En primer lugar, los hábitos son clave a la hora de determinar el coste médico de un cliente. En la primera gráfica se puede observar como el tabaco, el alcohol y la obesidad son tres claros factores de riesgo que incrementan mucho el coste médico. Incluso para exfumadores o fumadores recurrentes el coste médico es de media mucho mayor que para no fumadores. Además, la obesidad, como se puede ver, aumenta significativamente este gasto. El alcohol, aunque es difícil de controlar puesto que es un hábito muy común, casi todo el mundo consume algo de alcohol semanalmente muestra que en los casos más extremos de costes muy altos siempre se da en gente consumidora. Puedo confirmar después de un

análisis más exhaustivo que estos son los 3 factores claves, otros como las horas de sueño y el estrés influyen, pero no de igual medida.

- 2- **Clusters:** En el segundo gráfico se puede observar un modelo de clusters creado basados en los hábitos de los clientes y su edad. A modo de resumen es útil para entender cómo separar grupos y ponerles precio de forma única basados en cuáles han sido los factores más determinantes.
- 3- **Enfermedades crónicas:** En el tercer gráfico se puede observar cómo las enfermedades crónicas son claves a la hora de predecir el coste. De media que el paciente posea una de estas suma 4000\$ al coste. Algunas de las más caras son el cáncer y los problemas de riñón.
- 4- **Frecuencia de visitas:** Por último, muestro un gráfico donde podéis observar la relación entre el coste anual y la frecuencia de visitas hospitalarias. Como podéis ver existe una relación clara entre estos, aunque siempre se pueden observar excepciones sobre todo en el primer grupo.

### Prioridades Clave

Recomendamos centrar la estrategia de suscripción y gestión en los tres factores más predictivos y accionables: Hábitos, Frecuencia de Visitas y Edad.

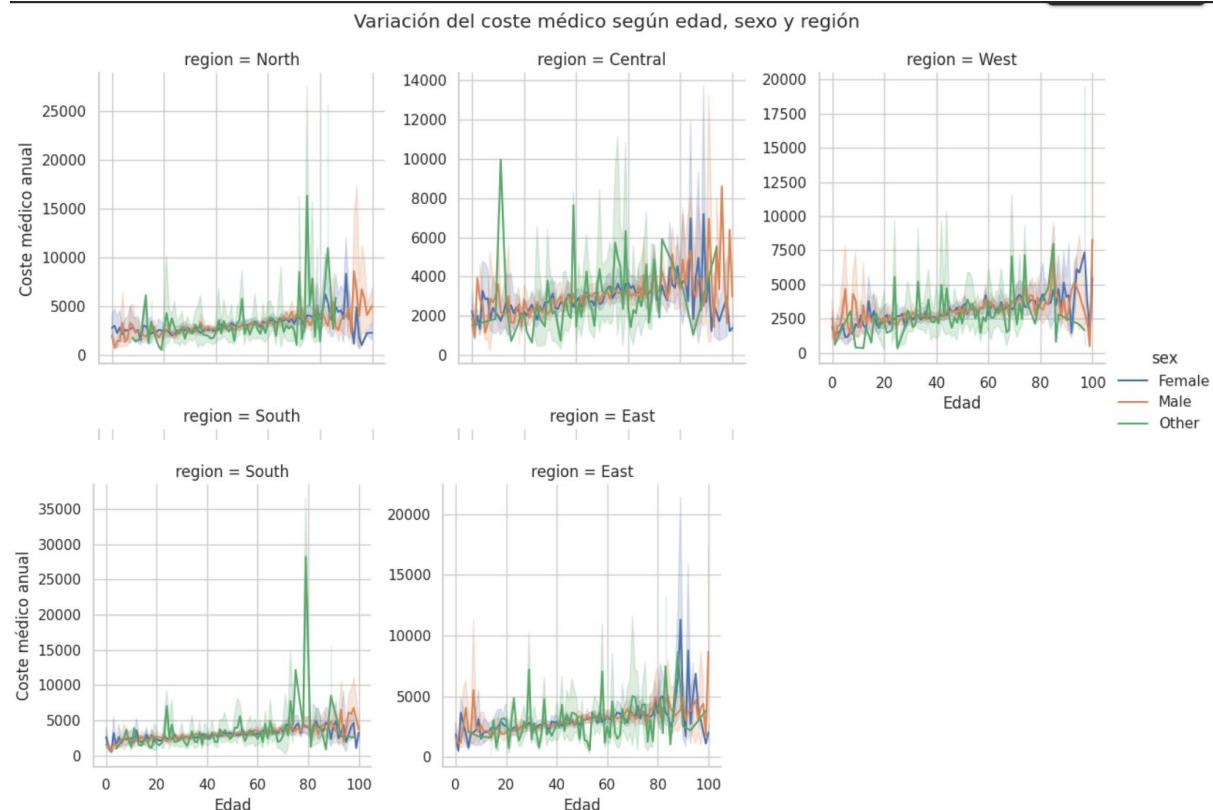
1. **Hábitos (Riesgo Controlable):** El tabaquismo, la obesidad y el consumo de alcohol son los principales impulsores de los costes. La aseguradora debe implementar urgentemente programas de bienestar con incentivos financieros directos para modificar estos comportamientos, ya que ofrecen el mayor retorno de la inversión en mitigación de riesgo.
2. **Modelos Predictivos:** Utilizar el modelo de clusters para asignar precios que reflejen fielmente el riesgo real del grupo (edad y hábitos). Esto garantiza una tarificación justa y la sostenibilidad de la cartera.
3. **Gestión de Cargas Crónicas:** Si bien la edad es un predictor base, el coste se dispara con la presencia de enfermedades crónicas y la alta frecuencia de visitas. Se requiere una gestión proactiva de casos para los pacientes crónicos y aquellos con alta utilización de servicios, minimizando complicaciones costosas.

La compañía debe pivotar hacia una tarificación predictiva y basada en el comportamiento individual. La inversión en la prevención de malos hábitos es más efectiva que cualquier ajuste en factores menos predictivos como el género o la educación, asegurando que los recursos se destinen a donde el riesgo es más palpable.

## 2. Gráficas del análisis exploratorio y breve explicación de cada una

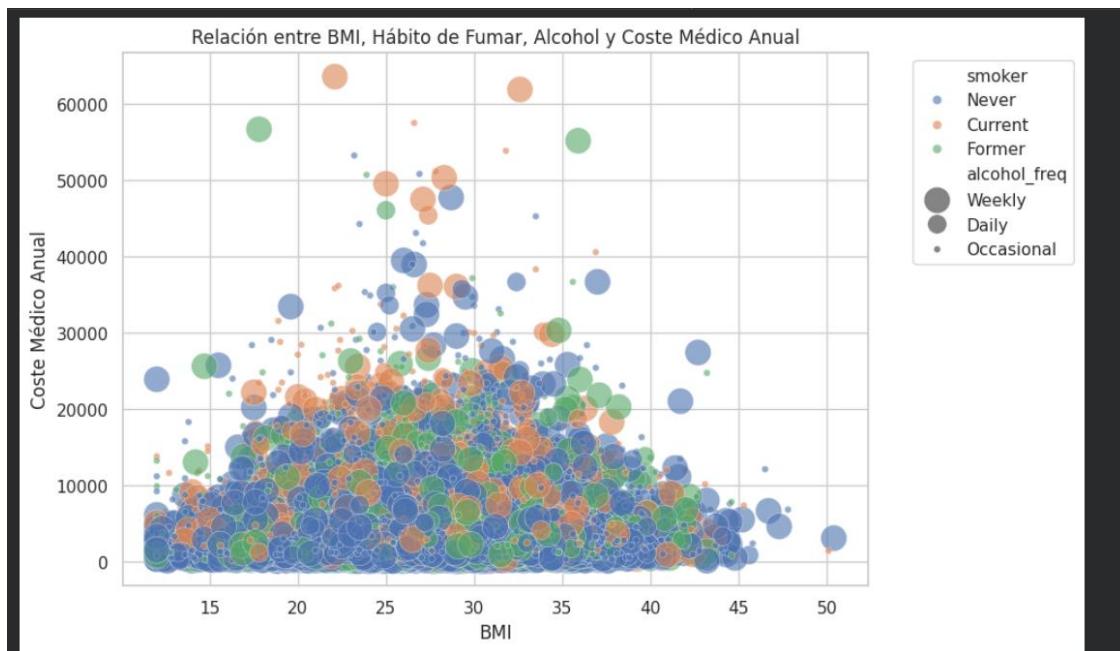
### Coste médico anual según región, sexo y edad:

De primeras vamos a intentar sacar relaciones que a grandes rasgos nos puedan ayudar a orientar el problema. Para ello vamos a relacionar las variables de sexo, edad y región para tratar de sacar conclusiones:



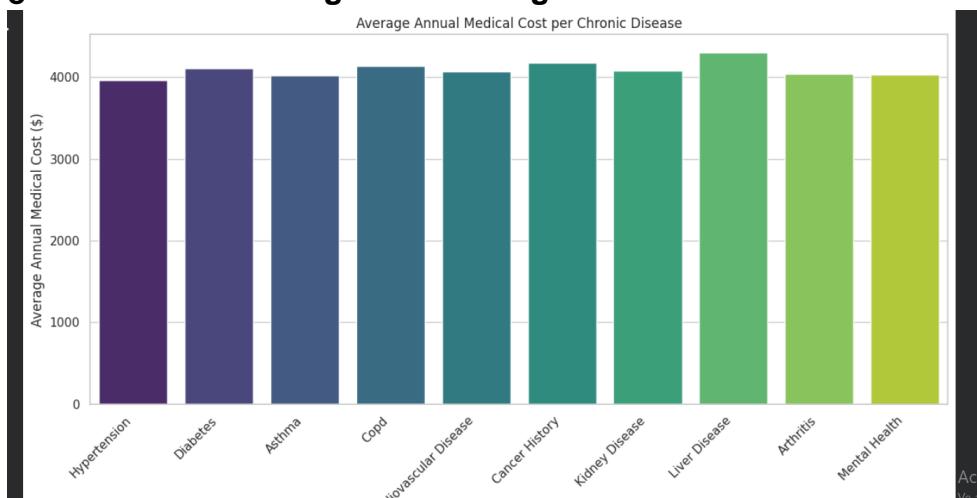
Se puede observar como obviamente la edad juega un papel clave en el gasto médica en todas las regiones y además la distribución es bastante similar en ellas. Si que existe la excepción de que en la región central hay valores atípicamente altos en las edades de los 20 y los 40 algo a tener en cuenta por la aseguradora. Además, en la región oeste el gasto médico es más bajo de media que en el resto. Por lo demás, en general a partir de los 75-80 años el gasto se dispara. En cuanto al sexo no observo grandes patrones, solo que las mujeres suelen vivir más y por lo tanto tienen más gasto a partir de los 90 porque los hombres se suelen morir antes. Ahora vamos a tratar de entender si los hábitos de vida juegan un papel importante en el gasto médico.

### Hábitos vs coste médico:



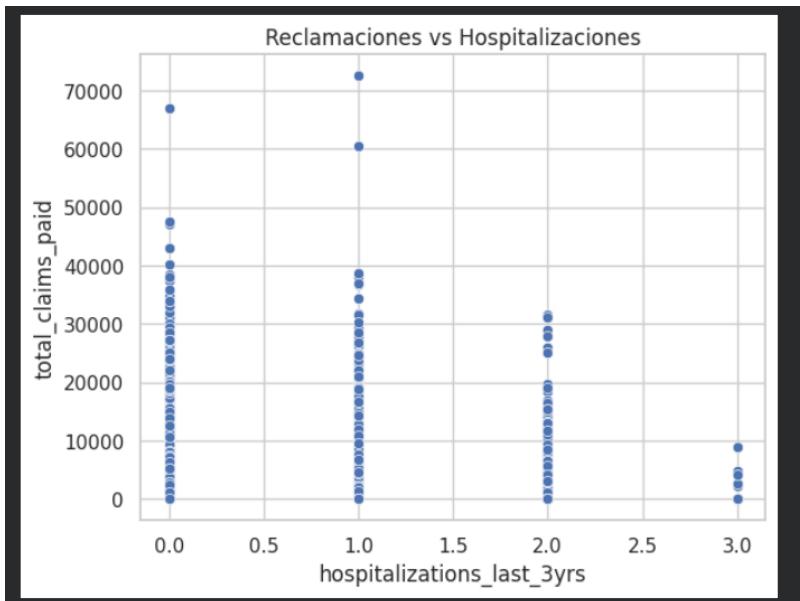
En cuanto a la relación del coste anual médico y los hábitos hay una clara tendencia entre la gente que fuma y la que no. La mayoría de fumadores tanto ocasionales como recurrentes e incluso los exfumadores tienen un gasto por encima de los 5000 que no se da tanto en no fumadores. Esto es clave a la hora de asegurar a esta gente. Por otra parte, el BMI nos grita algo obvio, cuanto más bajo menos gasto, aunque es verdad que a partir de los 30 puntos la relación se rompe, quizás porque el índice de mortalidad es mayor. El alcohol consumido semanalmente tambien influye sobretodo en casos más extremos.

### ¿Qué enfermedades generan más gasto?



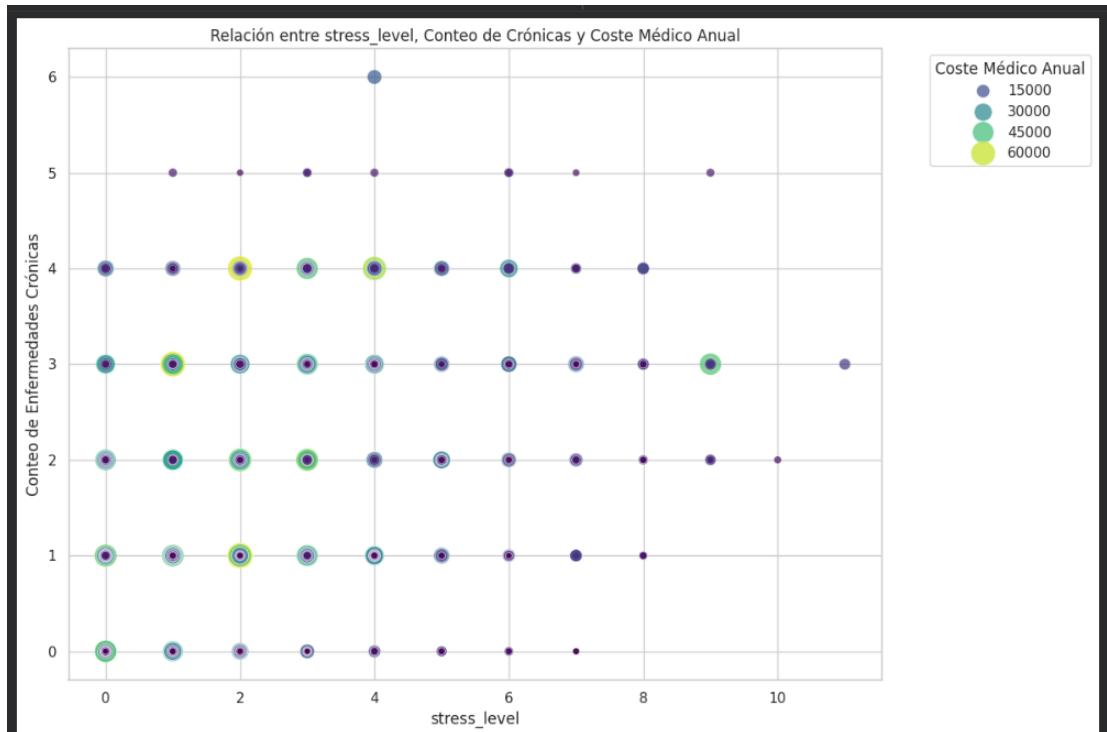
Las enfermedades crónicas son muy caras para una aseguradora por eso es importante identificar el gasto. Como se puede observar los problemas de riñon son los mas caros al igual que el cáncer. En general la mayoría de estas enfermedades producen gastos medios de unos 4000\$ anuales.

### ¿Hay relación entre las reclamaciones y la hospitalización previa del paciente?



Una de las posibles preguntas que cabe hacernos es si lo que se le reclama al seguro está relacionado con las hospitalizaciones. Curiosamente esto no es tan así. En general muchas de las reclamaciones no van relacionadas con que el cliente haya estado hospitalizado anteriormente.

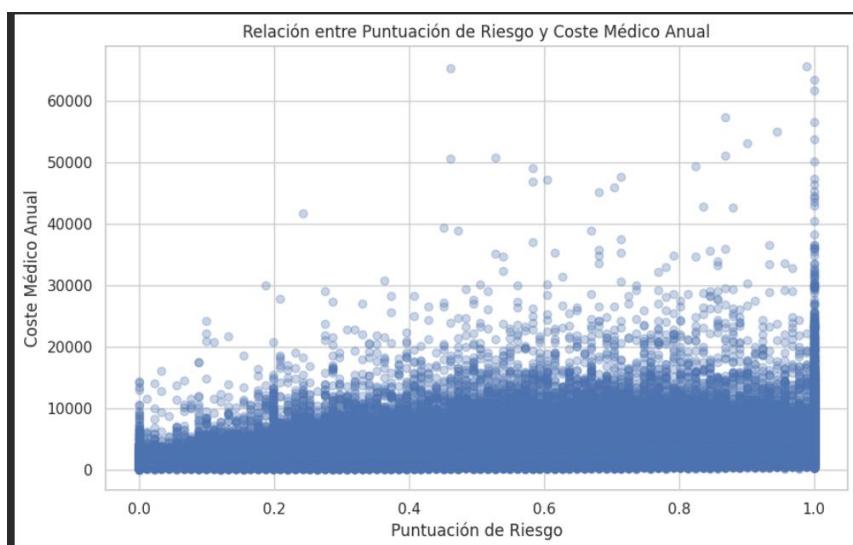
## ¿Qué otras variables son relevantes?



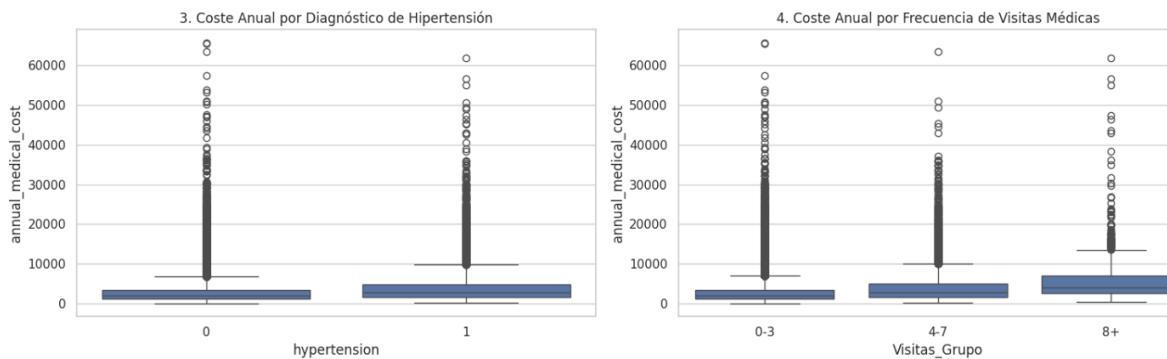
De cara a mi modelo de clustering considero importante analizar otras variables que puedan ser claves. Por ejemplo, en la gráfica superior se puede ver como el numero de enfermedades crónicas influye mucho en el coste médico anual. Por otro lado, aunque si afecta, el nivel de estrés no parece un factor tan determinante como otros.

## ¿La puntuación de riesgo es un buen indicador?

Por último, quiero comprobar si la puntuación de riesgo es un buen indicador del coste médico anual. Como se puede observar la relación es clara, aunque mejorable. A partir de 0.8 parece que el indicador falla un poco a la hora de determinar el coste médico.

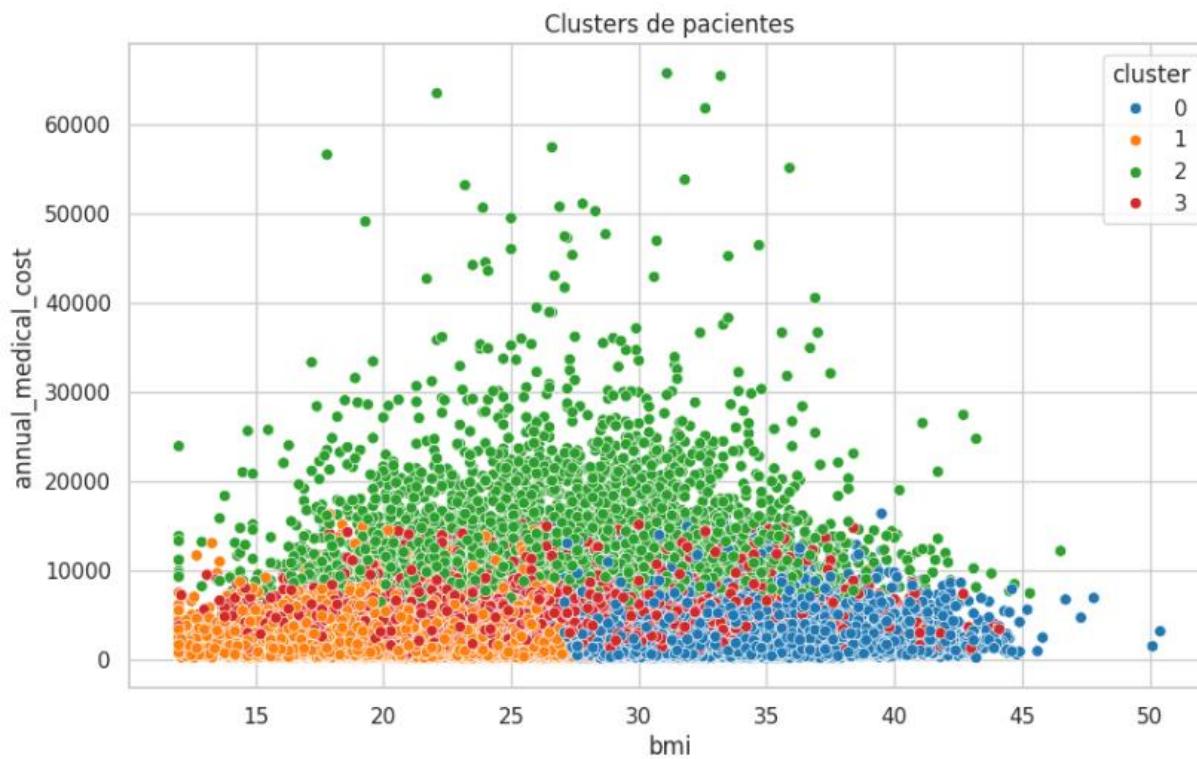


Otras gráficas interesantes muestran por ejemplo el coste por hipertensión y el coste por frecuencia de visitas médicas



### K-Means con conclusiones:

Ahora realizamos un modelo K-means con las conclusiones obtenidas del análisis exploratorio. Los malos hábitos tienen una gran influencia, la edad es un buen indicador de coste médico y las enfermedades crónicas son buenas predictoras de riesgo. Además las visitas claramente influyen en el coste:

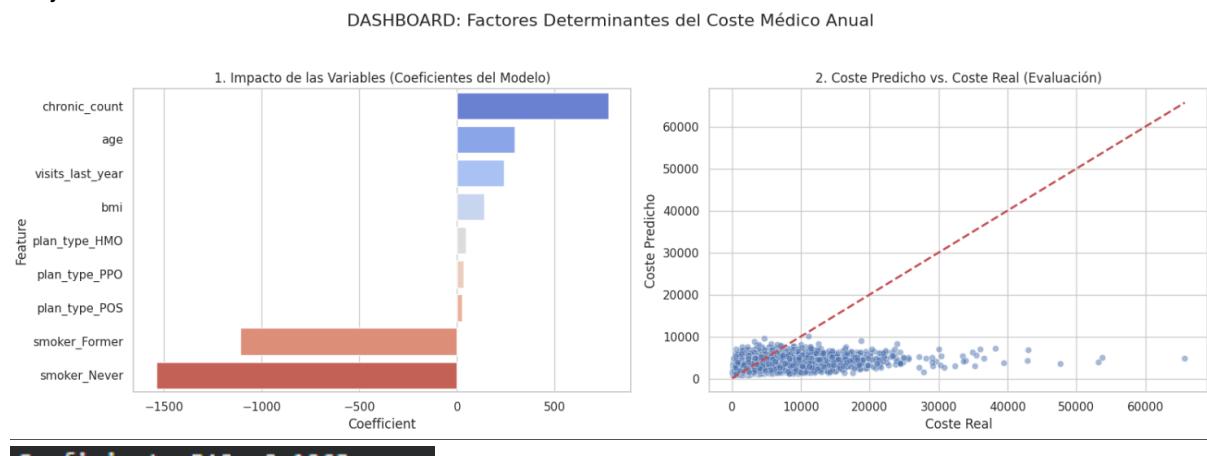


Aunque es una primera desagregación simple de primeras podemos observar tres grupos muy diferenciados.

### 3. Modelo predictivo explicado y con tablas

Por último, voy a realizar un modelo de regresión para tratar de predecir el coste médico de un cliente basandomé en los factores clave que he ido sacando. Tras algunos intentos he

añadido las visitas médicas como un factor importante ya que ayuda al modelo a predecir mejor.



Estos son los resultados obtenidos. Se puede ver como chronic count es la variable más importante al igual que la edad y las visitas médicas. En este caso el ser fumador no influye tanto. Aunque el modelo no predice demasiado bien, es útil para entender que variables realmente influyen.