

¿Que tan probable es la asistencia de cada paciente a su cita médica?

BOULCHOUK ALEXIS ADRIAN
Coderhouse – Data Scientist II / Comisión 90455



Índice

► Healthcare Appointment Dataset

I.	Contexto y Audiencia	3
II.	Preguntas de interés e Hipótesis	4
III.	Metadata	5
IV.	Análisis Exploratorio	6
V.	Conclusión final	10

Contexto y Audiencia

► Healthcare Appointment Dataset

- Contexto:

Todo científico de datos debe tener habilidades tanto matemáticas y lógicas como interpersonales. Hoy busco cumplir con dichos requerimientos para poder perfeccionarme en mi carrera como científico de datos.

El dataset que escogí en esta ocasión cuenta con datos relacionados a un centro de atención sanitaria de Brasil (**Healthcare Appointment**) y mi objetivo es poder definir qué tan probable es que los pacientes asistan a sus citas médicas programadas analizando datos sobre su salud y comportamientos habituales.

- Audiencia:

Cualquier tipo de clínica que cuente con datos referentes a la salud y comportamiento de sus pacientes, ya que de este modo lograría ayudar con la organización de la clínica, reduciendo los gastos y optimizando la asignación de turnos.

Preguntas de interés e Hipótesis

► Healthcare Appointment Dataset

- ¿Los pacientes más jóvenes poseen una asistencia menor debido a su grado de responsabilidad?
Hipótesis: Entendemos que sí, por el comportamiento normal de la sociedad, el porcentaje de asistencia debería verse fuertemente afectado por el grupo etario debido a la responsabilidad del mismo.
- ¿Mientras más grande es el tiempo de antelación con el que se pacta la cita, más probabilidad hay de que no asistan a la misma?
Hipótesis: Sí, hay mucha probabilidad de que no asistan al olvidarse de la misma, pasado más de 2 meses.
- ¿Al no pagar por la consulta médica, los pacientes con subsidio tienden a no asistir a las mismas?
Hipótesis: Los pacientes deberían tender a no asistir, ya que no generan un compromiso real con la cita al no tener un impacto monetario en sus finanzas.
- ¿Qué tan probable es que un paciente no asista a su cita médica al no recibir un SMS recordatorio?
Hipótesis: Es muy probable, ya que al no recibir el SMS confirmatorio, los pacientes tienden a olvidarse de la misma.
- ¿Cuáles son las variables más importantes a la hora de determinar si un paciente asistirá a su cita médica?
Hipótesis: Las variables más importantes deberían ser las que corresponden a las enfermedades y a los tiempos de antelación con los que se programa la cita médica.

Metadata

► Healthcare Appointment Dataset

Citas analizadas

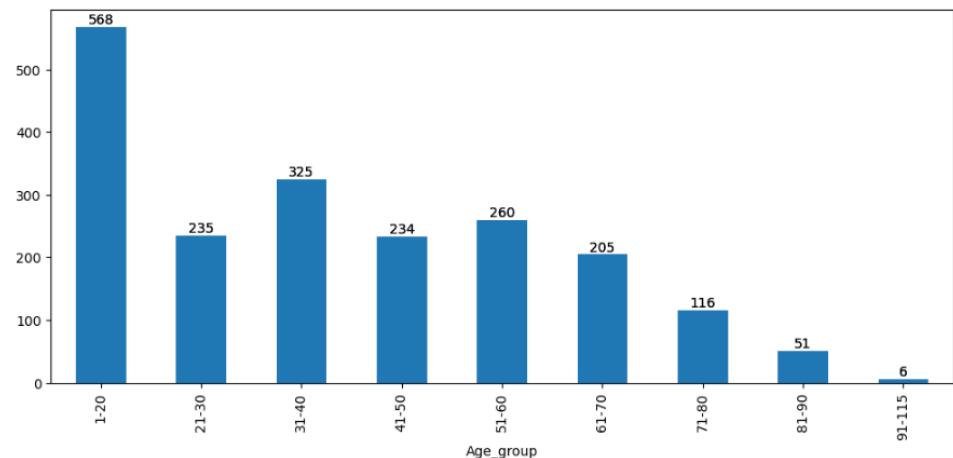


Puntos clave sobre salud y comportamiento diario

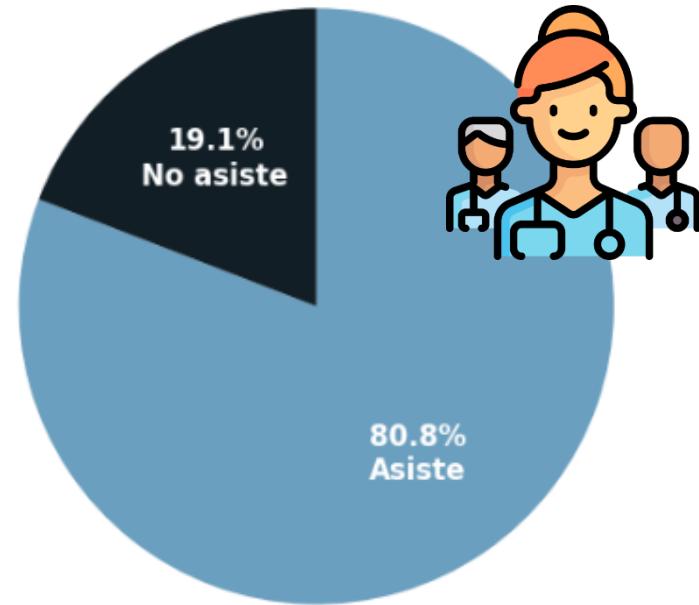
- Alcoholismo
- Hipertensión
- Discapacidad
- Diabetes



Distribución de la muestra por edad:



Porcentaje de asistencia

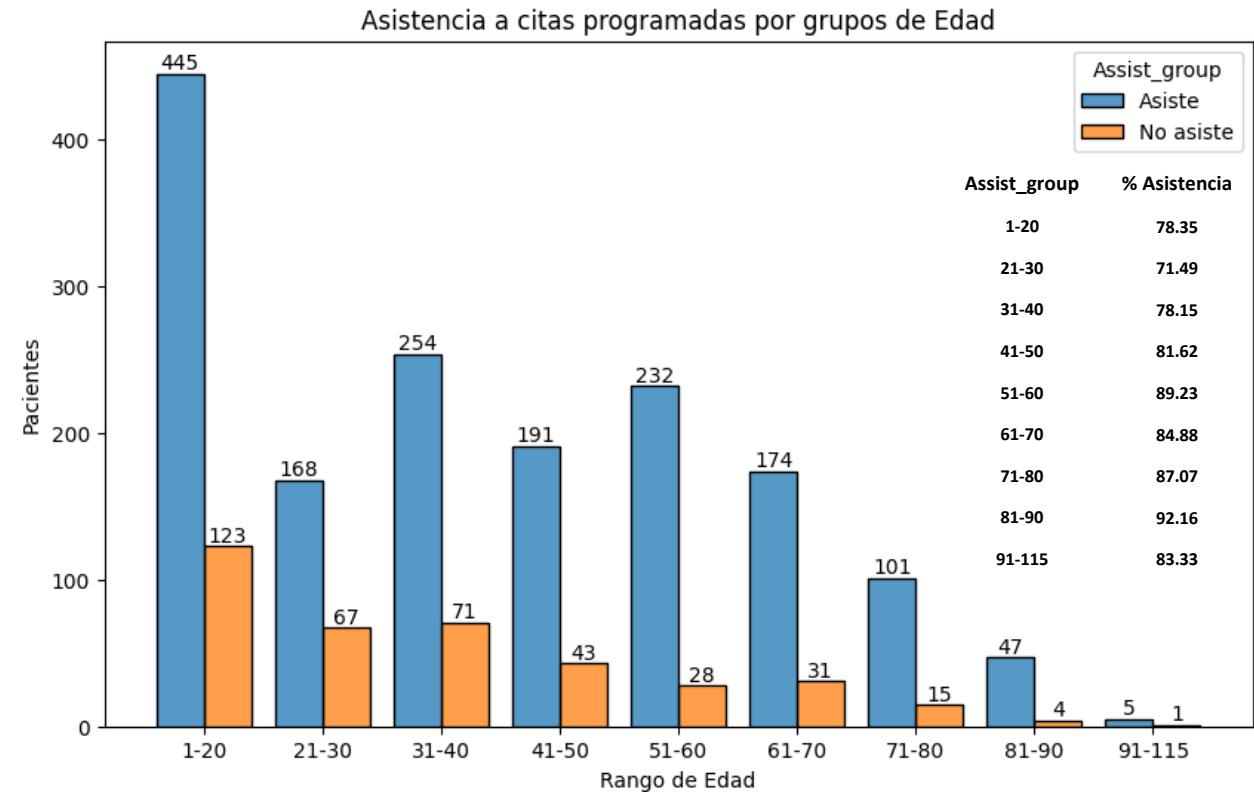


Análisis exploratorio

- ▶ ¿Los pacientes más jóvenes poseen una asistencia menor debido a su grado de responsabilidad?

Tomando en cuenta la siguiente clasificación: Niñez/Adolescencia (01-20) Juventud (21-30) Juventud adulta (31-40) Aduldez temprana (41-50) Aduldez intermedia (51-60) Aduldez plena (61-70) Aduldez mayor (71-80) Vejez inicial (81-90) Vejez avanzada (91-100)

Junto a los gráficos y tablas realizadas, notamos que efectivamente el grupo etario de "JUVENTUD" posee un porcentaje de asistencia menor (71.49%) al resto de los grupos, ya que entendemos que su grado de responsabilidad es el menor fuera del seno familiar (quitando de lado la niñez/adolescencia). A la vez, notamos que el % de asistencia va aumentando a medida que los pacientes se vuelven más grandes; tiene una caída leve comenzando la aduldez plena y vuelve a aumentar en la aduldez mayor, dandonos a entender que nuevamente poseen un tutor que ayuda a seguir su tratamiento. Ya en el grupo final, la asistencia vuelve a caer, mostrándonos que nuevamente prefieren no asistir a las citas.

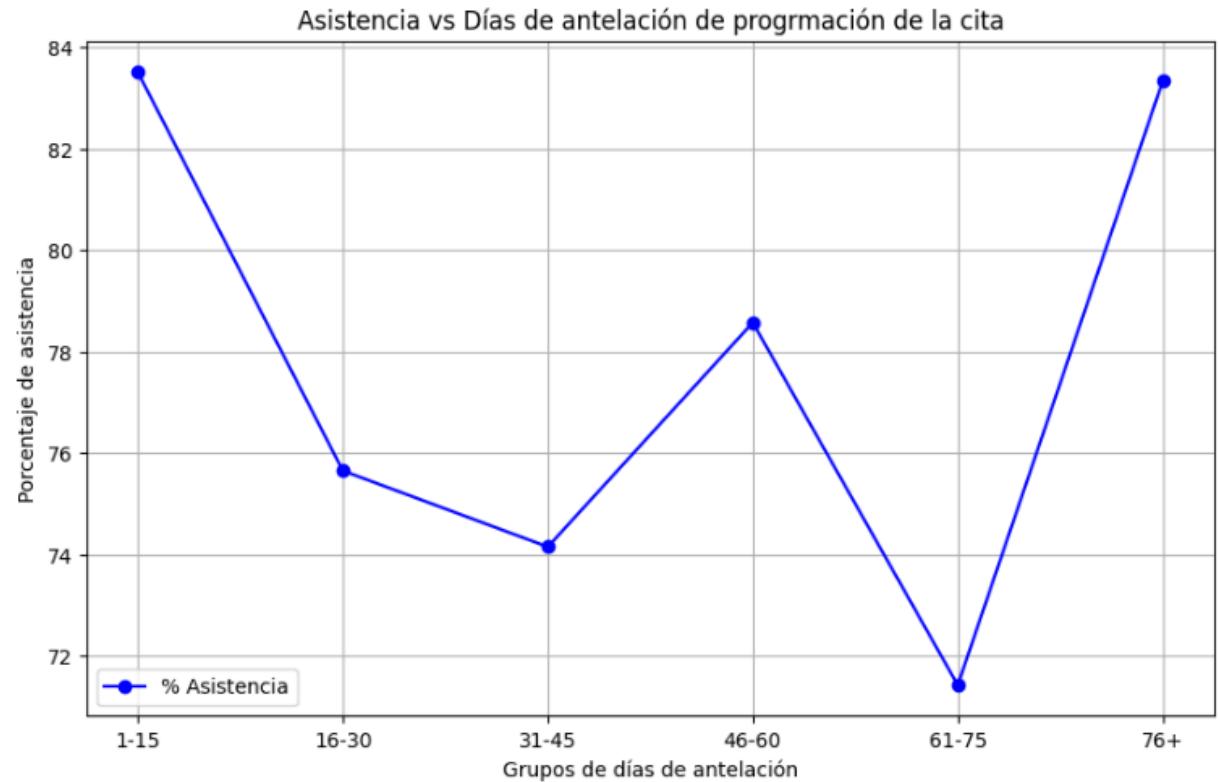


Análisis exploratorio

- ▶ ¿Mientras más grande es el tiempo de antelación con el que se pacta la cita, más probabilidad hay de que no asistan a la misma?

Analizando de manera gráfica y mediante tablas vemos que los días de antelación y las asistencias no son directamente proporcionales ya que cada punto posee porcentajes muy variados habiendo picos extremadamente altos en los extremos de los grupos por quincena y un pico alto dentro del grupo (46-60).

Como conclusión podemos decir que el porcentaje de asistencia es significativamente mayor cuando la programación de la cita es mas próxima o muy lejana a la fecha en la que se pide. El indicador no lleva una linea descendente constante a medida que la programación se hace mas lejana.

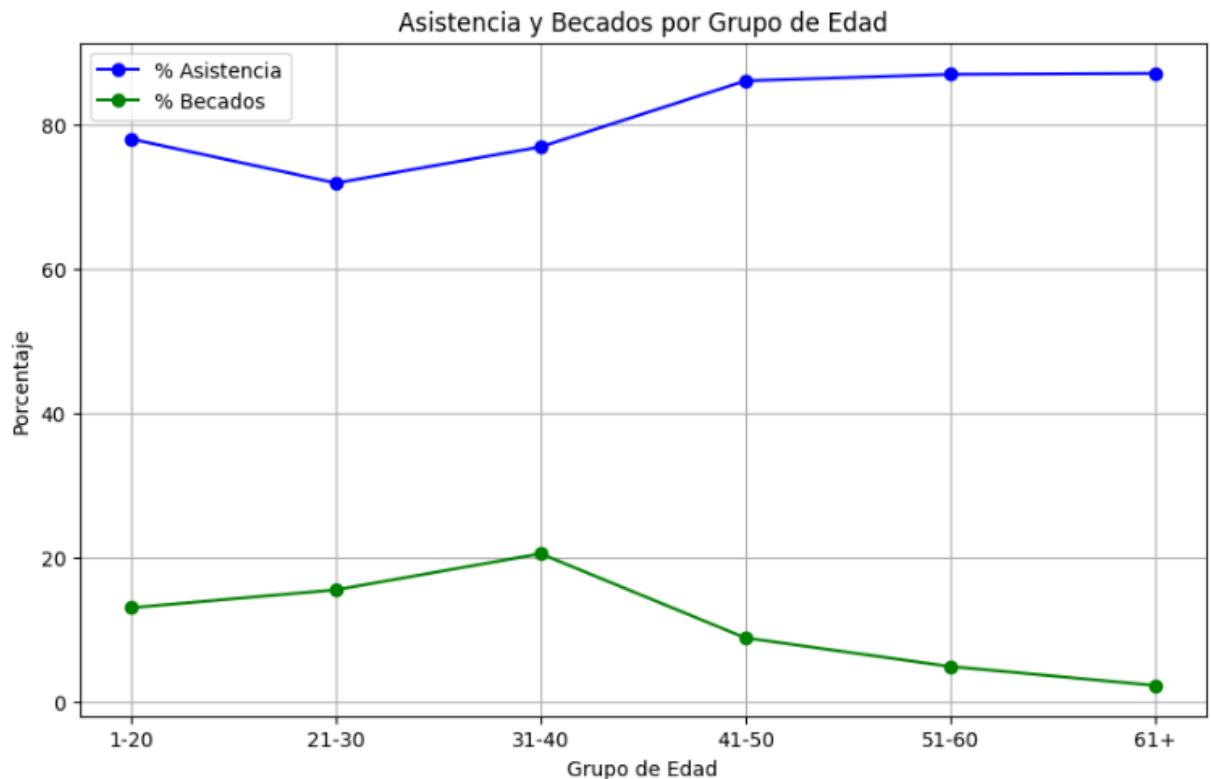


Análisis exploratorio

- ▶ ¿Al no pagar por la consulta médica, los pacientes con subsidio; tienden a no asistir a las mismas?

Analizando los datos por grupos etarios mediante gráficos y tablas, podemos notar que mientras más pacientes becados hay por grupo etario, menor es la cantidad de asistencia a las citas médicas, casi de manera proporcional.

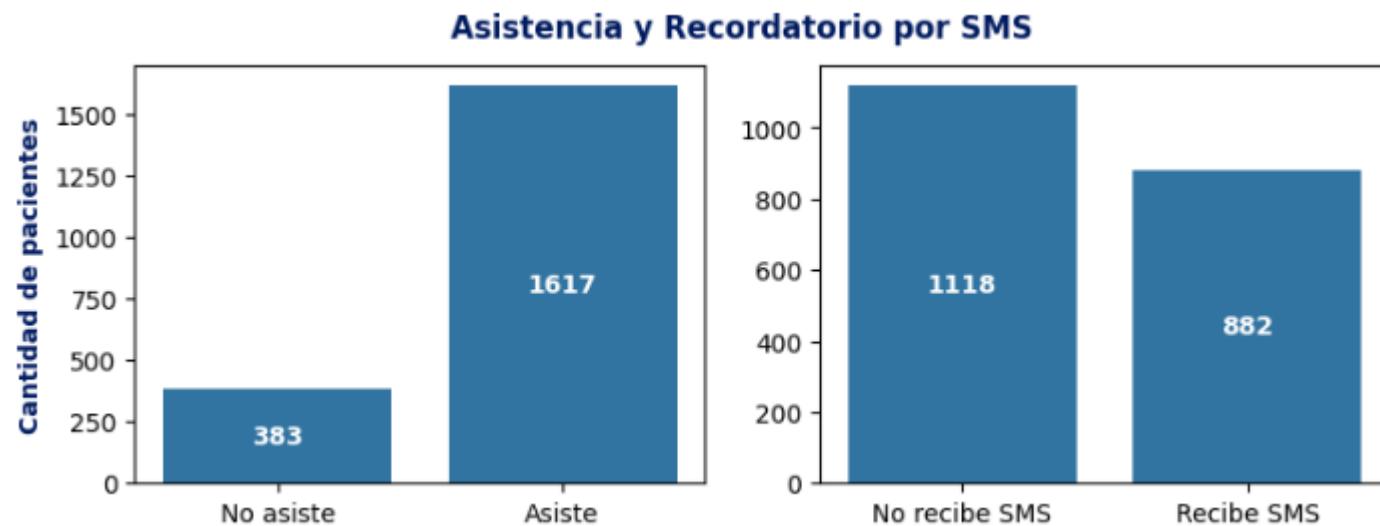
Sería necesario analizar más casos para poder verificar si se trata de una relación directamente proporcional o simplemente los pacientes más adultos poseen menos becas por factores externos.



Análisis exploratorio

- ▶ ¿Qué tan probable es que un paciente no asista a su cita médica al no recibir un SMS recordatorio?

Analizando los datos, distinto a la hipótesis planteada, vemos que el porcentaje de probabilidad de no asistencia al no recibir el SMS confirmatorio es más bajo que recibiéndolo (13.86% y 25.85% respectivamente), dándonos a entender que no es una variable directamente proporcional a la asistencia. Es probable que tengan otros motivos para no asistir a la cita programada.



Conclusión Final

- ▶ ¿Cuáles son las variables más importantes a la hora de determinar si un paciente asistirá a su cita médica?

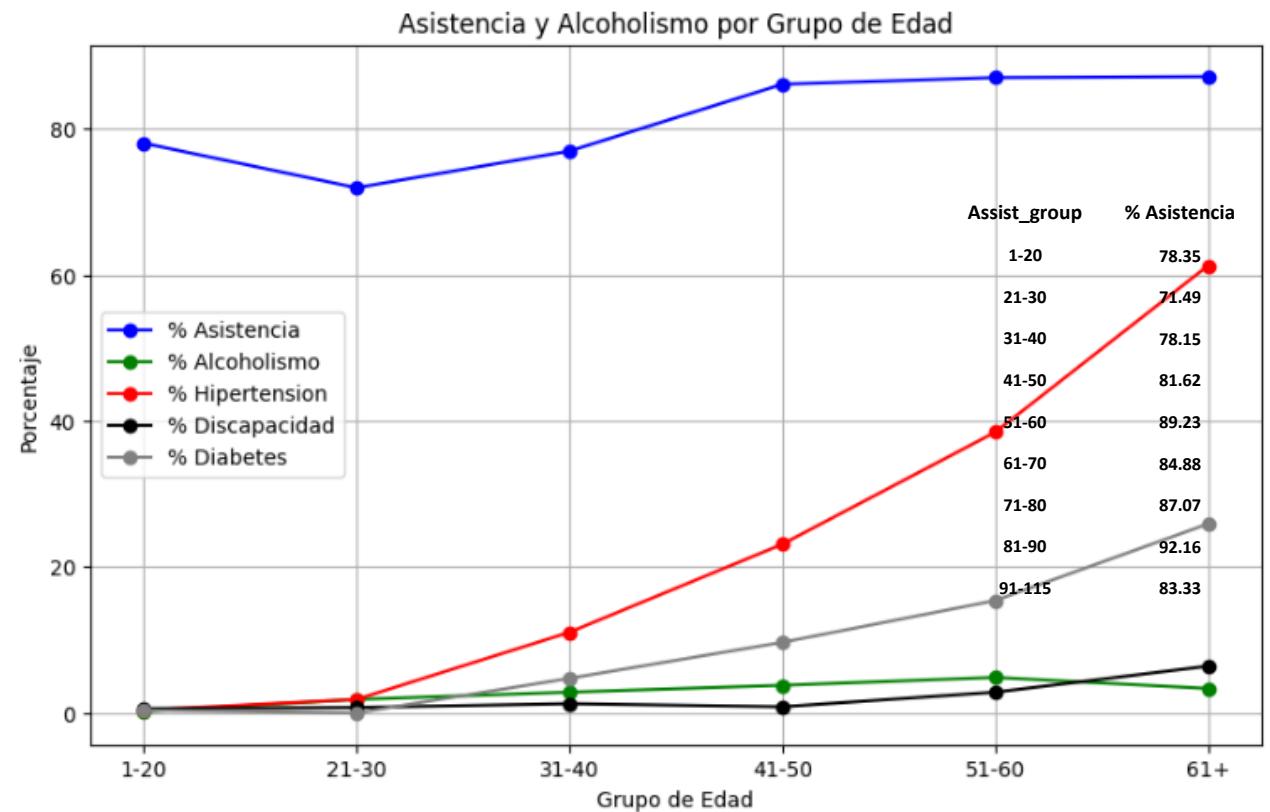
Analizando el gráfico de la derecha, podemos ver que, por el comportamiento de las líneas, no hay una relación directa entre el alcoholismo, la hipertensión, la discapacidad y la diabetes con la asistencia a las citas programadas. A su vez, vemos que estas variables sí están relacionadas entre sí y junto a la edad.

Conclusión preliminar (A confirmar una vez se realice el aprendizaje supervisado o no supervisado): Luego de todo el análisis realizado sobre las features, verificamos que las verdaderamente importantes corresponden a:

La edad. (Posee una fuerte relación debido a su CV alto).

Las becas. (Es casi directamente proporcional a la asistencia).

Los tiempos de antelación. (Posee una fuerte relación debido a su CV alto).



¡Gracias!

BOULCHOUK ALEXIS ADRIAN
Coderhouse – Data Scientist II / Comisión 90455

