📋 Executive Summary: Medical Appointment No-Show Analysis (April–June 2016)

This project investigates patient appointment adherence using a large dataset of medical appointments collected in Brazil from April to June 2016. With over 110,000 records, the dataset includes patient demographics, chronic conditions, appointment scheduling details, and whether the patient showed up. The central focus of the analysis is to understand **why patients miss appointments**, and what patterns or interventions (like SMS reminders) might help reduce no-shows.

The analysis begins with a general overview of appointment attendance. Out of all scheduled appointments, approximately **21% resulted in no-shows**, revealing a significant gap in care delivery. Interestingly, **over 42% of patients had more than one appointment**, underscoring the importance of studying **repeat patient behavior**, which may differ from that of one-time visitors.

Looking at monthly trends, **May recorded the highest no-show rate (~21.5%)**, compared to April (~18.4%) and June (~19.6%). This spike in May could be due to operational overload, public holidays, or seasonal variations in behavior. In terms of scheduling logistics, patients who waited longer between their scheduling and appointment dates were more likely to miss—especially in May. This was even more pronounced among patients with **chronic conditions like diabetes and hypertension**, who tended to face **longer average wait times** and **higher no-show rates**.

When evaluating the hour of the scheduled appointment, we found that no-show rates were lowest during early morning hours **(typically 7:00 AM to 9:00 AM)** and began to increase steadily as the day progressed. The highest no-show rates were observed **between 2:00 PM and 5:00 PM,** particularly among patients with **chronic conditions** such as diabetes and hypertension. Additionally, **young patients particularly males aged 10 to 29 consistently showed the highest no-show rates** across all months. On the other hand, older adults (60+) demonstrated the most reliable attendance, possibly due to a higher perceived need for care or more predictable routines.

The study also evaluated the effectiveness of **SMS reminders**. Contrary to expectations, patients who received an SMS reminder were actually **more likely to miss their appointment** in several groups. This may be explained by selection bias where SMS reminders were perhaps sent to high-risk patients or it could indicate that reminders alone are insufficient to drive behavioral change. This ineffectiveness was especially apparent among **frequent patients (6+ visits)**, where no-show rates remained above 25% even when SMS was received. This finding suggests that **more engaging interventions** such as two-way communication or personalized follow-ups might be necessary.

Next, the analysis explored how **patient frequency (visit count)** affects no-shows. Patients with more appointments were not only more likely to miss, but they also contributed **disproportionately to the total number of no-shows**. For instance, those with 6 or more visits represented a small fraction of patients but accounted for the **majority of missed appointments**. When stratifying by both visit frequency and chronic disease status, the **highest risk group emerged as patients with both diabetes and hypertension who had 6+ appointments** clearly indicating the need for special attention.

To consolidate these findings, a **logistic regression model** was built to predict the likelihood of a patient missing an appointment. The model included variables such as age, gender, wait time, SMS reminder status, chronic conditions, and total number of appointments. The results showed that:

- **Younger age** significantly increased the odds of missing an appointment.
- **Receiving an SMS** was associated with higher no-show likelihood—again supporting the earlier observation that reminders may not be effective without behavioral engagement.
- **Longer wait times** between scheduling and appointment date strongly predicted no-shows.
- **Diabetes** was a significant predictor of missed appointments, whereas hypertension was not.
- Interestingly, patients with more appointments had a slightly lower no-show risk, suggesting a marginal benefit from consistent engagement.

Overall, this analysis demonstrates that **missed appointments are not random**—they are patterned by age, condition, scheduling delay, communication method, and visit frequency. While broad solutions like SMS reminders help operationally, they are not sufficient on their own. Targeted, **data-driven strategies** particularly for younger, chronically ill, and high-frequency patients are necessary to meaningfully reduce no-shows and improve care delivery outcomes.