Project Proposal

In this project, the dataset that I am going to use is called **Medical Appointment No Shows**, was collected from **Kaggle** and can be found:

https://www.kaggle.com/joniarroba/noshowappointments

What I am looking for in this analysis has turned into questions which are as follows:

what is the reason that a patient/s will not show up for the appointment?

Is there any relationship with the age of the patient and his/her neighborhood that will affect the presence of the patient on the day of the appointment as well as making other to attend?

This model will help the healthcare sector finds out what possible reason to prevent with the aid of supplied features.

By providing such insights, the employee can make some adjustments to their appointments system to get better results as to predict whether a patient to show up or not in the day of the appointment.

The dataset consists of 13 features that will be used to answer my questions which are as follows:

- Gender
- DataMarcacaoConsulta: The day of the actual appointment, when they have to visit the doctor.
- DataAgendamento: The day someone called or registered the appointment, this is before appointment of course.
- Age
- Neighborhood
- Scholarship: in education and health more info https://en.wikipedia.org/wiki/Bolsa_Fam%C3%ADlia
- Hipertension
- Diabetes
- Alcoholism
- Handicap
- SMS_received

The target is to answer the question as to whether a patient to show up or not on appointment day.

Some exploratory data analysis and statistical information using Python libraries.

What would a minimum viable product (MVP) look like for this project?

Some statistical information and initial visualization.