**Hospitalization Hypothesis Testing**

Download Datasets

*This data project has been used as a take-home assignment in the recruitment process for the data science positions at Apollo Hospitals.*

Apollo Hospitals was established in 1983, renowned as the architect of modern healthcare in India. As the nation's first corporate hospital, Apollo Hospitals is acclaimed for pioneering the private healthcare revolution in the country.

**Assignment**

As a data scientist working at Apollo, the ultimate goal is to tease out meaningful and actionable insights from Patient-level collected data. You can help Apollo hospitals to be more efficient, influence diagnostic and treatment processes, and to map the spread of a pandemic.

One of the best examples of data scientists making a meaningful difference at a global level is in the response to the COVID-19 pandemic, where they have improved information collection, provided ongoing and accurate estimates of infection spread and health system demand, and assessed the effectiveness of government policies.

**The company wants to know:**

* Which variables are significant in predicting the reason for hospitalization for different regions;
* How well some variables like viral load, smoking, and severity level describe the hospitalization charges;

**Data Description**

The file **apollo\_data.csv** contains anonymized data of COVID-19 hospital patients and contains the following variables:

* **age** - an integer indicating the age of the primary beneficiary (excluding those above 64 years, since they are generally covered by the government)
* **sex** - the policy holder's gender, either male or female
* **smoker** - 'yes' or 'no' depending on whether the insured regularly smokes tobacco
* **region** - beneficiary's place of residence in Delhi, divided into four geographic regions - northeast, southeast, southwest, or northwest
* **viral load** - the amount of virus in an infected person's blood
* **severity level** - an integer indicating how severe the patient is
* **hospitalization charges** - individual medical costs billed to health insurance