Link: <https://www.kaggle.com/datasets/akashnath29/lung-cancer-dataset>

**About Dataset**

**Lung Cancer Dataset**

**Introduction:**

Lung cancer remains one of the most prevalent and deadly forms of cancer worldwide, posing significant challenges for early detection and effective treatment. To contribute to the global effort in understanding and combating this disease, we are excited to introduce our comprehensive Lung Cancer Dataset, now available on Kaggle.

**Scientific Overview:**

This dataset is an invaluable asset in the realm of Health Care, providing a structured foundation for the development of cancer detection models. This dataset exemplifies the variety of symptoms of Lung Cancer. Each category within the dataset—**'GENDER', 'AGE', 'SMOKING', 'YELLOW\_FINGERS', 'ANXIETY', 'PEER\_PRESSURE', 'CHRONIC\_DISEASE', 'FATIGUE', 'ALLERGY', 'WHEEZING', 'ALCOHOL\_CONSUMING', 'COUGHING', 'SHORTNESS\_OF\_BREATH', 'SWALLOWING\_DIFFICULTY', 'CHEST\_PAIN'**—has been carefully curated to encompass a diverse range of symptoms, ensuring that the resulting models are versatile and accurate. This scientific approach not only enhances the dataset's diversity to record symptoms of lung cancer but also contributes to the broader field of AI-driven health technologies, pushing the boundaries of what health care assistants can achieve.

**Dataset Composition**

The Lung Cancer Dataset includes a diverse array of symptoms essential for comprehensive analysis and model development. The primary categories of data are as follows:

**1. Patient Demographics**

Age: Provides the age at diagnosis, enabling analysis of age-related incidence and outcomes.  
Gender: Includes information on patient gender, facilitating gender-based studies.  
Smoking Status: Categorized as current smoker, former smoker, or non-smoker, this data is critical for evaluating the impact of smoking on lung cancer risk and progression.

**2. Medical History**

Comorbidities: Details additional health issues such as chronic obstructive pulmonary disease (COPD), which are relevant for treatment planning and prognosis.

**3. Clinical Data**

Vital Signs: Records of blood pressure, heart rate, respiratory rate, and other vital signs at diagnosis and during treatment.