**Objective – Explore the dataset and extract insights from the data. Using statistical evidence to**

* Prove (or disprove) that the medical claims made by the people who smoke is greater than those who don't?
* Prove (or disprove) with statistical evidence that the BMI of females is different from that of males.
* Is the proportion of smokers significantly different across different regions?
* Is the mean BMI of women with no children, one child, and two children the same?

Context[¶](https://www.kaggle.com/code/yogidsba/insurance-claims-eda-hypothesis-testing#Context)

Leveraging customer information is of paramount importance for most businesses. In the case of an insurance company, attributes of customers like the ones mentioned below can be crucial in making business decisions.

**Data Set**

* Age :- This is an integer indicating the age of the primary beneficiary (excluding those above 64 years, since they are generally covered by the government).
* Sex :- This is the policy holder's gender, either male or female.
* BMI :- This is the body mass index (BMI), which provides a sense of how over or under-weight a person is relative to their height. BMI is equal to weight (in kilograms) divided by height (in meters) squared. An ideal BMI is within the range of 18.5 to 24.9.
* Children :- This is an integer indicating the number of children / dependents covered by the insurance plan.
* Smoker :- This is yes or no depending on whether the insured regularly smokes tobacco.
* Region :- This is the beneficiary's place of residence in the U.S., divided into four geographic regions - northeast, southeast, southwest, or northwest.
* Charges​ :- Individual medical costs billed to health insurance

**Question to be answered**

* Are there more Male beneficary ?
* Are there more smoker ?
* Which region has maximum , medical cost billed to health insurance.?
* What is age of beneficary.?
* Do beneficary having more dependents had more medical cost billed.?