

## Laboratorio 4

### Información del dataset

En el archivo titulado **insurance.csv** se pueden encontrar datos generales de distintas personas y cuánto cobra una compañía de seguro por asegurarlas. El dataset está compuesto por las columnas age, sex, bmi, children, smoker, region y charges.

Siendo los siguientes los tipos correspondientes para dichas columnas:

1. age de tipo int64
2. sex de tipo int64
3. bmi de tipo float64
4. children de tipo int64
5. smoker de tipo int64
6. region de tipo int64
7. charges de tipo float64

### Hipótesis u objetivo

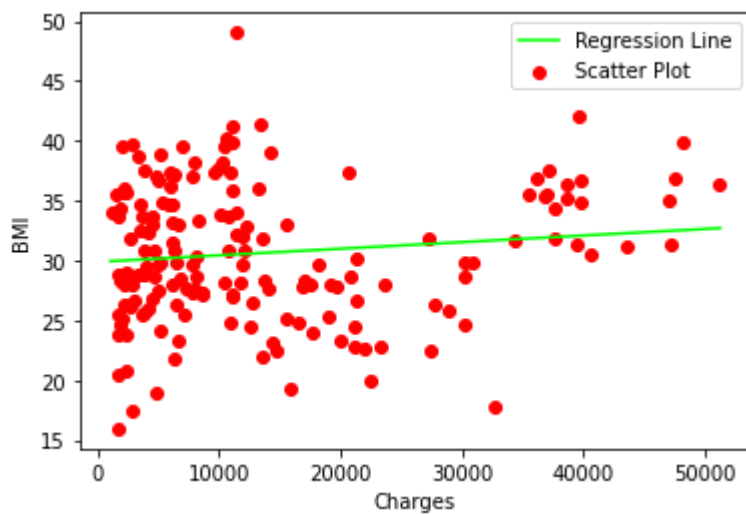
Generar una regresión lineal y/o polinomial que nos permita predecir el costo de seguro de una persona de acuerdo a su perfil.

### Solución y exploración:

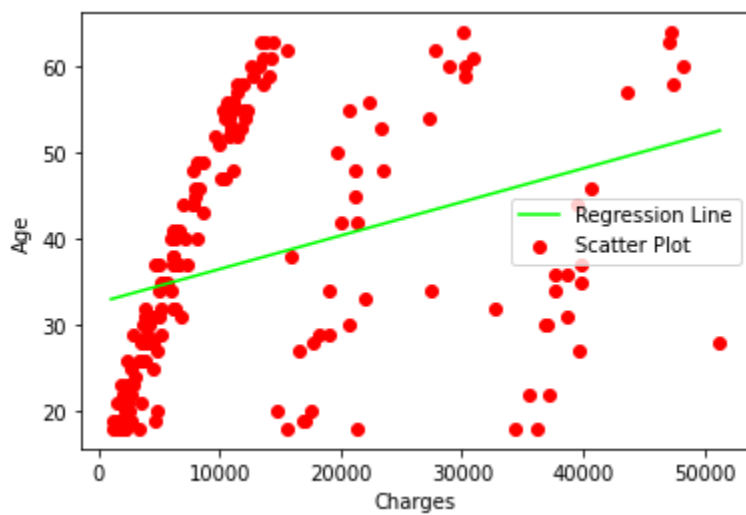
Primero se hizo una limpieza y preparación de los datos para asegurarnos de que no hubieran datos en el dataset que nos pudieran afectar los resultados. Luego separamos nuestro dataset en dos grupos: El primero siendo data de prueba y el segundo siendo data de entrenamiento. Finalmente realizamos los modelos de regresión lineal y polinomial comparando las medidas MAE, MSE y RMSE.

### Resultados (Se encuentran en el documento notebook.ipynb)

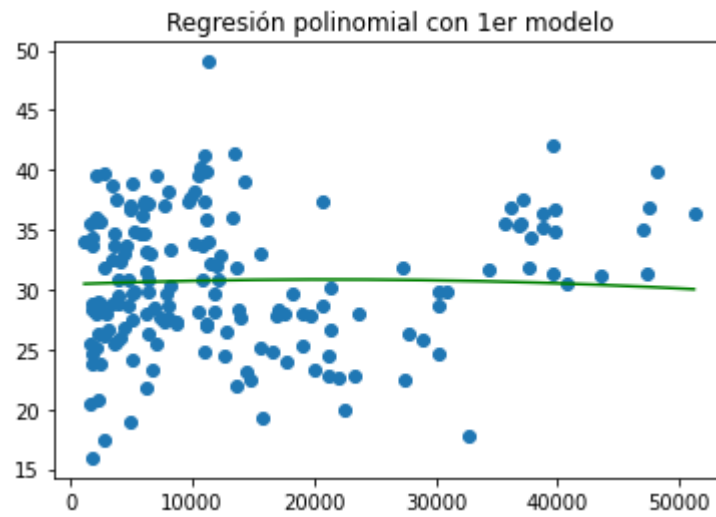
### Regresión Lineal (cargos vs bmi)



### Regresión Lineal (cargos vs edad)



### Regresión polinomial (cargos vs bmi)



### Regresión polinomial (cargos vs edad)

