



SEP
SECRETARÍA DE
EDUCACIÓN PÚBLICA



TECNOLÓGICO NACIONAL DE MÉXICO



INSTITUTO TECNOLÓGICO DE TIJUANA

CARRERA

INGENIERÍA EN SISTEMAS COMPUTACIONALES

MATERIA
MINERÍA DE DATOS

TAREA

PRÁCTICA 1, UNIDAD #3

FECHA ENTREGA
17/05/2022

ALUMNO(A)

HOWARD HERRERA ERWIN #18210716
PÉREZ LÓPEZ ALICIA GUADALUPE #18210514

DOCENTE

JOSE CHRISTIAN ROMERO HERNANDEZ



Importaremos el archivo a R

```
# Importing the
dataset <- read.csv(file.choose())
dataset <- read.csv('Salary_Data.csv')
```

Vamos a instalar las librerías, usaremos una nueva

```
# splitting the dataset into the Training set and Test set
# Install.packages('caTools')
install.packages("caTools")
library(caTools)
library(ggplot2)
```

Vamos a crear una serie de números con los sig comandos.

```
library(ggplot2)
set.seed(123)
split <- sample.split(dataset$salary, SplitRatio = 2/3)
training_set <- subset(dataset, split == TRUE)
test_set <- subset(dataset, split == FALSE)
```

La regresión del modelo lm o lineal con la relación entre salario y años de experiencia.

Variable de predicción de datos:

```
# Fitting Simple Linear Regression to the Training set
regressor = lm(formula = Salary ~ YearsExperience,
               data = dataset)
summary(regressor)

# Predicting the Test set results
y_pred = predict(regressor, newdata = test_set)
```

Ahora con nuestra librería de ggplot2 correremos las siguientes líneas, estas nos ayudaran a darle color a la gráfica, nombre y tamaño

```
# Visualising the Training set results
library(ggplot2)
ggplot() +
  geom_point(aes(x=training_set$YearsExperience, y=training_set$Salary), color = 'red') +
  geom_line(aes(x = training_set$YearsExperience, y = predict(regressor, newdata = training_set)),color = 'blue') +
  ggtitle('Salary vs Experience (Training Set)') + xlab('Years of experience') + ylab('Salary')

# Visualising the Test set results
ggplot() +
  geom_point(aes(x=test_set$YearsExperience, y=test_set$Salary),color = 'red') +
  geom_line(aes(x = training_set$YearsExperience, y = predict(regressor, newdata = training_set)),color = 'blue') + ggtitle('Salary vs Experience (Test
```



Muestra de gráficas.

