**a) Importar el archivo de datos empleados.xls**

**Datos/Importar datos/desde un archivo de Excel**

empleados <- readXL("C:/Desktop/empleados.xls", rownames=FALSE, header=TRUE, na="", sheet="Respuestas", stringsAsFactors=TRUE)

**b) Realizar el diagrama de dispersión del peso en función de la edad**

**Gráficas/Diagramas de dispersión**

**scatterplot(Peso~Edad, reg.line=FALSE, smooth=FALSE, spread=FALSE, boxplots=FALSE,**

 **span=0.5, ellipse=FALSE, levels=c(.5, .9), data=empleados)**

**c) Ajustar el modelo de regresión lineal peso en función de la edad. Representar la recta ajustada**

**Estadísticos/Ajuste de modelos/Regresión lineal**

RegLineal <- lm(Peso~Edad, data=empleados)

summary(RegLineal)

Contraste de regresión

$$\left\{\begin{array}{c}H\_{0}: Reg=0\\H\_{1}: Reg\ne 0\end{array}\right. ≡ \left\{\begin{array}{c}H\_{0}: β\_{1}=0\\H\_{1}: β\_{1}\ne 0\end{array}\right.$$

En este caso obtenemos un valor del estadístico $F$ de 6.765 con un p\_valor = 0.01075. Por lo tanto rechazamos la hipótesis nula, y se concluye que existe relación lineal entre las variables

Contrastes sobre coeficientes de regresión individuales

$$\left\{\begin{array}{c}H\_{0}: β\_{0}=0 \\H\_{1}: β\_{0}\ne 0 \end{array}\right.\left\{\begin{array}{c}H\_{0}: β\_{1}=0 \\H\_{1}: β\_{1}\ne 0 \end{array}\right.$$

Los p\_valores de la constante (1.93e-14), y la edad (0.0108). Por lo que rechazamos las hipótesis nulas planteadas.

La recta ajustada aparece especificada a través de sus dos coeficientes

$$Peso= 58.2563+ 0.8043\*Edad$$

El error estándar de los residuos tiene un valor de 12.75

Respecto a la bondad del ajuste, el coeficiente de determinación $R^{2}$ tiene un valor de 0.06519 indica que el 6.51% de toda la variabilidad que tiene el fenómeno relativo al peso puede ser explicado por el logaritmo de la edad.

El coeficiente de correlación lineal 0.05556. Por lo tanto, la dependencia es positiva, pero muy débil.

**Gráficas/Diagramas de dispersión/Opciones**

scatterplot(Peso~Edad, reg.line=lm, smooth=FALSE, spread=FALSE, boxplots=FALSE,

 span=0.5, ellipse=FALSE, levels=c(.5, .9), data=empleados)

**d) Realizar una regresión logarítmica considerando que el peso de los empleados depende de la edad**

**Estadísticos/Ajuste de modelos/Modelo lineal**

> RegresionLogaritmica <- lm(Peso ~ log(Edad), data=empleados)
> summary(RegresionLogaritmica)

Contraste de regresión

$$\left\{\begin{array}{c}H\_{0}: Reg=0\\H\_{1}: Reg\ne 0\end{array}\right. ≡ \left\{\begin{array}{c}H\_{0}: β\_{1}=0\\H\_{1}: β\_{1}\ne 0\end{array}\right.$$

En este caso obtenemos un valor del estadístico $F$ de 9.533 con un p\_valor = 0.002632. Por lo tanto rechazamos la hipótesis nula, y se concluye que existe relación lineal entre las variables

Contrastes sobre coeficientes de regresión individuales

$$\left\{\begin{array}{c}H\_{0}: β\_{0}=0 \\H\_{1}: β\_{0}\ne 0 \end{array}\right.\left\{\begin{array}{c}H\_{0}: β\_{1}=0 \\H\_{1}: β\_{1}\ne 0 \end{array}\right.$$

Los p\_valores de la constante (0.99902), y la edad (0.00263) indican que no podemos rechazar la hipótesis nula planteada respecto a la constante.

La recta ajustada aparece especificada a través de sus dos coeficientes

$$Peso= -0.02984+ 24.87238\*ln⁡(Edad)$$

El error estándar de los residuos tiene un valor de 12.58

Respecto a la bondad del ajuste, el coeficiente de determinación $R^{2}$ tiene un valor de 0.08948 indica que el 8.94% de toda la variabilidad que tiene el fenómeno relativo al peso puede ser explicado por el logaritmo de la edad.

El coeficiente de correlación lineal 0.08009. Por lo tanto, la dependencia es positiva, pero muy débil.

**e) Realiza el gráfico con la función scatterplot**

**scatterplot(Peso~log(Edad),reg.line=lm, xlab="Edad", ylab="Peso", data=empleados)**

**f) Realizar el modelo exponencial y comparar ambos modelos**

**Estadísticos/Ajuste de modelos/Modelo lineal**



> RegresionExponencial <- lm(log(Peso) ~ Edad, data=empleados)
> summary(RegresionExponencial)

Contraste de regresión

$$\left\{\begin{array}{c}H\_{0}: Reg=0\\H\_{1}: Reg\ne 0\end{array}\right. ≡ \left\{\begin{array}{c}H\_{0}: β\_{1}=0\\H\_{1}: β\_{1}\ne 0\end{array}\right.$$

En este caso obtenemos un valor del estadístico $F$ de 6.727 con un p\_valor = 0.01097. Por lo tanto rechazamos la hipótesis nula, y se concluye que existe relación lineal entre las variables

Contrastes sobre coeficientes de regresión individuales

$$\left\{\begin{array}{c}H\_{0}: β\_{0}=0 \\H\_{1}: β\_{0}\ne 0 \end{array}\right.\left\{\begin{array}{c}H\_{0}: β\_{1}=0 \\H\_{1}: β\_{1}\ne 0 \end{array}\right.$$

Los p\_valores de la constante (2e-16), y la edad (0.010335 indican que rechazamos las hipótesis nulas planteadas.

La recta ajustada aparece especificada a través de sus dos coeficientes

$$ln⁡(Peso)= 4.087817+ 0.010335\*Edad$$

El error estándar de los residuos tiene un valor de 0.1642

Respecto a la bondad del ajuste, el coeficiente de determinación $R^{2}$ tiene un valor de 0.06486 indica que el 6.48% de toda la variabilidad que tiene el fenómeno relativo al logaritmo del peso puede ser explicado por la edad.

El coeficiente de correlación lineal 0.05521. Por lo tanto, la dependencia es positiva, pero muy débil.

Ambos modelos los comparamos mediante el coeficiente de determinación

Coeficiente de determinación del modelo exponencial = 0.06486

 Coeficiente de determinación del modelo logarítmico = 0.08948

Por lo tanto el modelo logarítmico explica una mayor variabilidad de los datos

g) **Ajustar el Peso en función de la Altura y el cuadrado de la Edad**



**Estadísticos/Ajuste de modelos/Modelo lineal**

****

> RegresionEdadPeso <- lm(Peso ~ Edad + I(Altura^2), data=empleados)
> summary(RegresionEdadPeso)

Contraste de regresión

$$\left\{\begin{array}{c}H\_{0}: Reg=0\\H\_{1}: Reg\ne 0\end{array}\right. ≡ \left\{\begin{array}{c}H\_{0}: β\_{1}=0\\H\_{1}: β\_{1}\ne 0\end{array}\right.$$

En este caso obtenemos un valor del estadístico $F$ de 24.05 con un p\_valor = 3.405e-9. Por lo tanto rechazamos la hipótesis nula, y se concluye que existe relación lineal entre las variables

Contrastes sobre coeficientes de regresión individuales

$$\left\{\begin{array}{c}H\_{0}: β\_{0}=0 \\H\_{1}: β\_{0}\ne 0 \end{array}\right.\left\{\begin{array}{c}H\_{0}: β\_{1}=0 \\H\_{1}: β\_{1}\ne 0 \end{array}\right.$$

Los p\_valores de la constante (0.3402), la edad (0.0142) y el cuadrado de la altura (1.28e-8) indican que no podemos rechazar la hipótesis nula referente a la constante.

La recta ajustada aparece especificada a través de sus dos coeficientes

$$Peso= -12.04+ 0.6580\*Altura-0.002335(Edad^{2})$$

El error estándar de los residuos tiene un valor de 10.82

Respecto a la bondad del ajuste, el coeficiente de determinación $R^{2}$ tiene un valor de 0.3338 indica que el 33.38% de toda la variabilidad que tiene el fenómeno relativo al peso puede ser explicado por la altura y el cuadrado de la edad.

El coeficiente de correlación lineal 0.32. Por lo tanto, la dependencia es positiva, pero débil.

Por lo tanto el modelo logarítmico explica una mayor variabilidad de los datos