**EJERCICIO DE REGRESIÓN LINEAL**

**Materia de Minería de datos Periodo escolar: 2022-1**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grupo: |  |  | Equipo: |  |

Nombre de los integrantes del equipo:

|  |  |
| --- | --- |
| 1) |  |
| 2) |  |
| 3) |  |

Las ventas de línea blanca varían dependiendo de la oferta de casas nuevas, cuando se incrementa la venta de casas nuevas también crece el de los electrodomésticos (lavaplatos, lavadoras de ropa, secadoras y refrigeradores). Una empresa compiló los datos que se muestran en la siguiente tabla.

Se desea crear un modelo matemático que represente la relación de los datos. Utilice la guía proporcionada.

Ejercicio adaptado de Levín, Rubín, Balderas, Del Valle y Gómez. (2004). Estadística para administración y economía. Séptima Edición. Prentice-Hall.

|  |  |
| --- | --- |
| **Construcción de casas (miles)** | **Venta de línea blanca (miles)** |
| 2.0 | 5.0 |
| 2.5 | 5.5 |
| 3.2 | 6.0 |
| 3.6 | 7.0 |
| 3.3 | 7.2 |
| 4.0 | 7.7 |
| 4.2 | 8.4 |
| 4.6 | 9.0 |
| 4.8 | 9.7 |
| 5.0 | 10.0 |

Responder cada uno de los siguientes incisos. Agregar la generación de tablas de cálculos y la presentación de las fórmulas que utilice en cada sección.

1. Generar la gráfica de variables
2. Realice los cálculos pasos a paso para generar la ecuación de regresión.
3. Realice la verificación de la ecuación de regresión de una recta generada con el método de mínimos cuadrados.
4. Realice los siguientes cálculos (muestre el proceso)

a) Suma de cuadrados debida al error

b) Suma total de cuadrados

c) Suma de cuadrados debida a la regresión

d) El coeficiente de determinación

e) Exprese el significado del coeficiente de determinación encontrado

f) El coeficiente de correlación y su significado

1. Calcule los errores estándar de la estimación
2. Los intervalos de confianza
3. Aplique la prueba t para determinar si el modelo es estadísticamente significativo
4. Genere la ecuación de recta en el Knime incorporando prueba de normalidad y gráfico de residuales.