Ejercicio Individual AE2

| Plan Formativo: Ciencia de Datos | Nivel de Dificultad |
| --- | --- |
| Módulo 3: Análisis exploratorio y programación estadística. | Bajo / medio |
| Tema: Correlación y Regresión. | |  | | --- | |
| Intención del aprendizaje o aprendizaje esperado: | |
| | * Explica los conceptos de correlación y regresión lineal para la caracterización de un conjunto de datos de una población. | | --- | | |
| Ejercicios planteados | |
| **Caso**:  Una compañía desea hacer predicciones del valor anual de sus ventas totales en cierto país a partir de la relación de éstas y la renta nacional. Para investigar la relación cuenta con los siguientes datos:   | **X** | **Y** | | --- | --- | | 189 | 402 | | 190 | 404 | | 208 | 412 | | 227 | 425 | | 239 | 429 | | 252 | 436 | | 257 | 440 | | 274 | 447 | | 293 | 458 | | 308 | 469 | | 316 | 469 |   X representa la renta nacional en millones de euros e Y representa las ventas de la compañía en miles de euros en el periodo que va desde 2010 hasta 2020 (ambos inclusive). Calcular:   1. La recta de regresión lineal de Y sobre X. 2. El coeficiente de correlación lineal e interpretarlo. 3. Si en 2011 la renta nacional del país fue de 325 millones de euros. ¿Cuál será la predicción para las ventas de la compañía en este año?   **Solución**   * La **recta de regresión** de Y sobre X  | **Xi** | **Yi** | **Xi2** | **Yi2** | **Xi \* Yi** | | --- | --- | --- | --- | --- | | 189 | 402 | 35721 | 161604 | 75978 | | 190 | 404 | 36100 | 163216 | 76760 | | 208 | 412 | 43264 | 169744 | 85696 | | 227 | 425 | 51529 | 180625 | 96475 | | 239 | 429 | 57121 | 184041 | 102531 | | 252 | 436 | 63504 | 190096 | 109872 | | 257 | 440 | 66049 | 193600 | 113080 | | 274 | 447 | 75076 | 199809 | 122478 | | 293 | 458 | 85849 | 209764 | 134194 | | 308 | 469 | 94864 | 219961 | 144452 | | 316 | 469 | 99856 | 219961 | 148204 | | **2753** | **4791** | **708933** | **2092421** | **1209720** |  * **El**coeficiente de correlación lineal**e interpretarlo**   Es un coeficiente de correlación positivo y cercano a uno, por lo que la correlación es directa y fuerte.   * Si en 2011 la renta nacional del país fue de 325 millones de euros. ¿Cuál será la predicción para las ventas de la compañía en este año? | |
| Datos de apoyo al planteamiento | |
| Tablas de frecuencia.  Gráficos y formas de distribución.  Medidas de tendencia.  Medidas de dispersión. | |
| Preguntas guía | |
| * ¿Qué es la estadística descriptiva? * ¿Qué son las tablas de frecuencia? * ¿Cuáles son las medidas de tendencia central? * ¿Qué son las medidas de dispersión? | |
| Recursos Bibliográficos: | |
| [1] Correlación  <https://www.incibe-cert.es/blog/correlacion-herramientas-analisis-datos>  [2] Diagrama de Dispersión y correlación lineal  <https://www.youtube.com/watch?v=rv0Xy8edFRg>  [3] Tablas de Contingencia  <https://www.youtube.com/watch?v=CEWJZxFpC8w>  <https://conceptosclaros.com/que-es-tabla-contingencias-par-que-sirve/>  [4] Coeficiente de Pearson  <https://www.uv.es/webgid/Descriptiva/31_coeficiente_de_pearson.html>  [5] Regresión Lineal  <https://www.youtube.com/watch?v=SsFBnvkoZa4> | |