| DATOS PERSONALES | FIRMA |
| --- | --- |
| |  |  | | --- | --- | | Nombre: Sindy Katerine | DNI:1032439654 | |  |  | |  |
| Apellidos: Rincón Torres |

| ESTUDIO | ASIGNATURA | CONVOCATORIA |
| --- | --- | --- |
| MÁSTER UNIVERSITARIO EN INGENIERÍA MATEMÁTICA Y COMPUTACIÓN (PLAN 2016) | 4391020006.- TÉCNICAS MULTIVARIANTES | Ordinaria Número periodo 1823 |

| FECHA | MODELO | CIUDAD DEL EXAMEN |
| --- | --- | --- |
| 14-16/01/2022 | Modelo - D | Bogotá |

| Etiqueta identificativa |
| --- |
|  |

**INSTRUCCIONES GENERALES**

1. Ten disponible tu documentación oficial para identificarte, en el caso de que se te solicite.
2. Si tu examen consta de una parte tipo test, indica las respuestas en la plantilla según las características de este.
3. Debes contestar en el documento adjunto, respetando en todo momento el espaciado indicado para cada pregunta. Si este es en formato digital, los márgenes, el interlineado, fuente y tamaño de letra vienen dados por defecto y no deben modificarse. En cualquier caso, asegúrate de que la presentación es suficientemente clara y legible. Entrega toda la documentación relativa al examen, revisando con detenimiento que los archivos o documentos son los correctos. El envío de archivos erróneos o un envío incompleto supondrá una calificación de “no presentado”.
4. Durante el examen y en la corrección por parte del docente, se aplicará el Reglamento de Evaluación Académica de UNIR que regula las consecuencias derivadas de las posibles irregularidades y prácticas académicas incorrectas con relación al plagio y uso inadecuado de materiales y recursos.
5. No está permitido el uso de Internet ni ningún tipo de comunicación con otra persona.Durante todo el examen tu teléfono móvil debe estar en modo avión.
6. La parte principal de cada pregunta consiste en interpretar y comentar los resultados obtenidos. Si te limitas a hacer los cálculos no vas a poder superar el examen.
7. Es fundamental que las respuestas estén debidamente redactadas, de forma clara y precisa y sin faltas de ortografía.
8. Para hacer el examen puedes utilizar los apuntes del curso y los scripts que hayas preparado y Python para hacer los cálculos.

**Puntuación**

**Preguntas**

* Puntuación máxima 10.00 puntos

El examen constará de un ejercicio práctico (8,5 puntos) y una pregunta teórica (1,5 puntos). Los enunciados están en la página 14 y el espacio para responder el examen está entre las práginas 4 y 13.

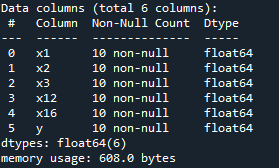
 (Responder en 10 caras)  
  
**1.** Pregunta

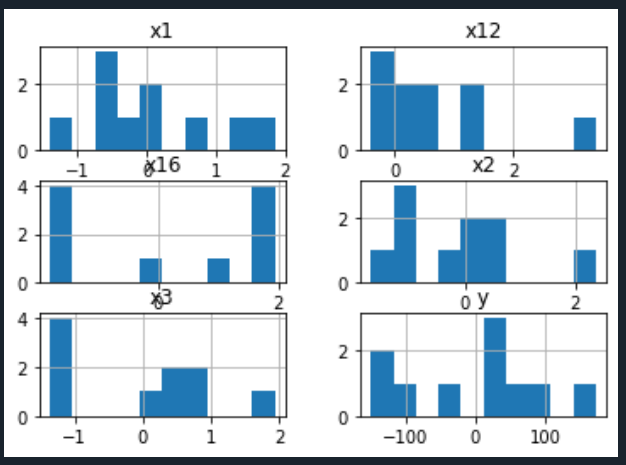
Fecha de nacimiento: 21/09/1990

M=0+9+3=12

d=16

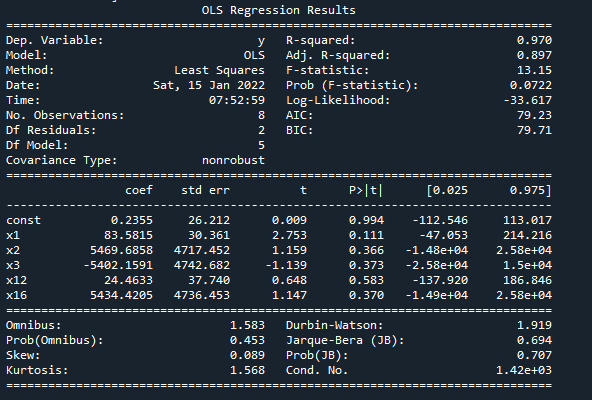
* En total son 10 datos para muestra en opinion es una muestra muy pequeña y se tienen 6 variables junto con la variable respuesta, en cuenta a su distribucion los tadota no son homogenios y no contienen una distribucion normal. Como se observa en las siguientes graficas





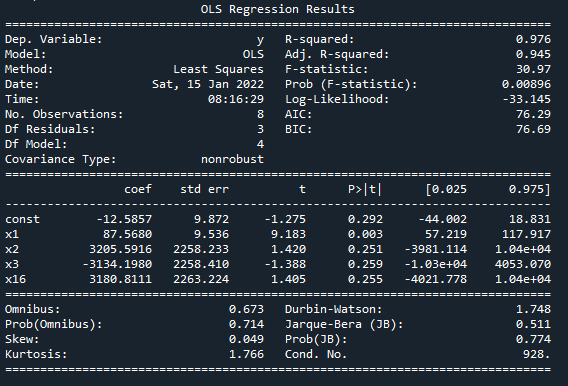
Se realiza la particion necesaria

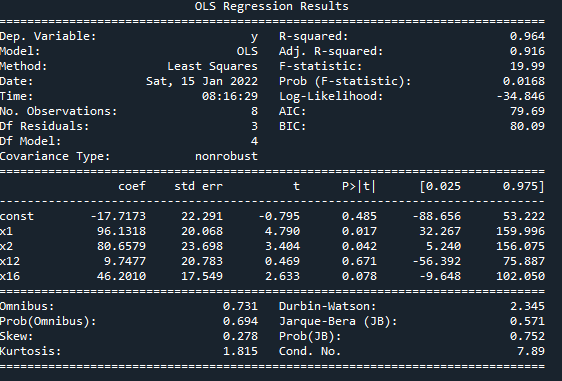
* REGRESION LINEAL MULTIVARIANTE



Se observa un de 0.97 y un r ajustado de 0.897 valores muy altos y una diferencia muy pequeña (0.073), una diferencia despresiable, sin embargo aunque el nivel de confianza es alto, al observar los P valor nos damos cuenta que las variables no tienen Correlación con la variable respuesta “y”, estas variables no son significativas, esto tal ves debido a la muestra tan pequeña que se tiene puesto que a tener mas registros podríamos tener un mejor análisis.

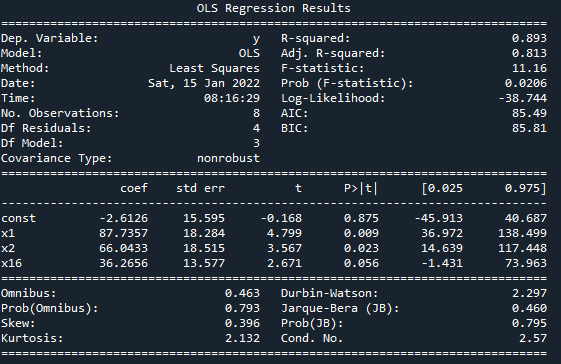
* Stepwise





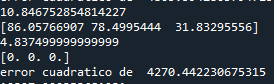
Despues de analizar el modelo de regression multiple eliminamos la variable con el p valor mas alto que para este caso seria “x12”, observando un de 0.97 y un r ajustado de 0.94, tambien se observa que la variable “x1” empieza a tener un valor mas pequeño teniendo correlacion con la respuesta y ajustando mejor el modelo, en cambio si solo eleminamos la variable “x3” tenemos  de 0.96 y un r ajustado de 0.91, tambien se observa que la variable “x1” y “x2”empieza a tener un valor mas pequeño teniendo correlacion con la respuesta y ajustando mejor el modelo, en cambio si solo eleminamos la variable.

Al eliminar estas dos variables “x12” y “x3” obtenemos un tenemos  de 0.86 y un r ajustado de 0.81, observamos que ya las variables “x1” y “x2” empiezan a tener una correlacion mas alta con la variable respuesta “y”, lo que significa que son variables significativas, acabe aclara que aunque aunque el modelo esta ajustado la muestra aun es muy pequeña y el error cuadratico es de 3185.768243195267, lo que indica que el ajuste no ha sido significativo.



* Lasso

Para lasso se toma = , aunque es un método que aplica penalización se observa la eliminación de variables y se observa que de igual forma el error cuadrático sigue siendo alto



Conclusion:

Al analizar la información resultante de cada uno de los métodos, no se escoge ninguno de los tres, pues ninguno tiene un error significante para el modelo , puestos los valores cuadraticos de los tres son demaciado altos , esto se debe tal ves a la muestra de datos tan pequeña y a los outlier de los datos



