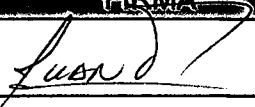


DATOS PERSONALES		FIRMA
Nombre: JUAN PABLO	DNI: 79953128	
Apellidos: POVEDA LOTERO		

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



02557866

Máster Universitario en Ingeniería Matemática y
Computación (Plan 2016) | 1823
4391020006- Técnicas Multivariantes | 1823




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.

Preguntas

- Puntuación máxima 10.00 puntos

NO UTILIZAR ESTA PARTE DE LA HOJA

DATOS PERSONALES		FIRMA
Nombre:	JUAN PABLO	
DNI:	79953128	
Apellidos:	POVEDA LOTERO	

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 páginas 4 y 13.

1. Pregunta

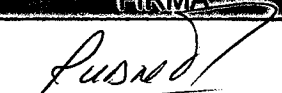
**NO UTILIZAR ESTA
PARTE DE LA HOJA**

Pregunta (1.5 puntos)

La presencia de datos anómalos en una muestra acarrea las siguientes consecuencias:

- Se reduce la confiabilidad de indicadores como las medidas de tendencia central: Media, mediana, moda, sus valores podrían diferir mucho.
- Se distorsiona el análisis de los datos al momento de realizar cálculos que requieran el uso de curvas de distribución de probabilidad.
- Al momento de tabular los datos para la creación de tablas o histogramas podría llevar a obtenerse intervalos en los cuales no hay datos. Esta sería entre otras, una forma de detectar esta situación.
- Teniendo en cuenta que valores como la media deben ser cercanos para la población y la muestra, ante la presencia de un valor anómalo en una de las muestras la diferencia entre estas será mayor sin importar la técnica mediante la cual se extrajo la muestra, de modo que, si por ejemplo se aplicó una técnica por conglomerado aquella muestra con el valor anómalo presentará una media significativamente diferente a las otras muestras.

Ante la posibilidad de trabajar con datos anómalos existe la opción de utilizar las herramientas de la estadística robusta, denominada así porque es menos sensible a estos datos.

DATOS PERSONALES		FIRMA
Nombre: JUAN PABLO	DNI: 79953128	
Apellidos: POVEDA LOTERO		

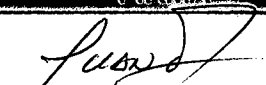
Pregunta (8.5 puntos)


Fecha de Nacimiento 30/10/79 $\Rightarrow d = 16$
 $m = 4$

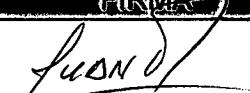
x_1 x_2 x_3 x_4 x_{16} y

ADRI

ESTADÍSTICA 120101

DATOS PERSONALES		FIRMA
Nombre:	JUAN PABLO	
DNI:	79953128	
Apellidos:	POVEDA LOTERO	

DATOS PERSONALES		FIRMA
Nombre:	JUAN PABLO	
DNI:	79953128	
Apellidos:	POVEDA LOTERO	

DATOS PERSONALES		FIRMA
Nombre:	JUAN PABLO DNI: 79953128	
Apellidos:	POVEDA LOTERO	

Preguntas - Pregunta 1

Ejercicio (8.5 puntos)

Escribe tu fecha de nacimiento (por ejemplo, 30/04/1987) y realiza los siguientes cálculos

- Calcula m como la suma de los dígitos del mes en que naciste y el número 3 (en el ejemplo anterior, $m = 0 + 4 + 3 = 7$)
- Calcula d en función del día que naciste. Si el día que naciste $\in [1, 10]$, entonces $d = 14$, si $\in [11, 20]$ entonces $d = 15$ y si por el contrario es ≥ 21 entonces $d = 16$ (en el ejemplo anterior, $30 \geq 21 \Rightarrow d = 16$)

Para hacer el examen debes considerar un *dataset* formado por las variables $x_1, x_2, x_3, x_m, x_d, y$. (en el ejemplo anterior, serían las variables $x_1, x_2, x_3, x_7, x_{16}, y$) de la Tabla ??

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	y
-0.53	-1.69	0.26	0.04	0.27	0.18	-0.15	-0.65	-0.52	-1.54	-2.91	1.37	0.25	-2.21	-0.02	1.94	-128.65
0.80	-1.24	0.48	0.38	0.55	-1.17	0.37	0.60	1.04	1.05	1.68	0.09	0.49	-0.45	1.76	1.72	24.66
0.01	-1.01	0.61	-0.81	-0.22	1.76	0.77	0.29	1.01	1.78	-1.32	0.51	-0.25	-1.00	1.23	1.62	13.61
1.39	-1.02	-1.20	-0.73	0.61	0.19	1.64	-0.10	0.54	1.80	0.23	0.53	-0.44	0.37	-1.01	-0.18	46.52
0.03	-0.01	1.93	-2.10	-0.82	-0.89	-2.84	0.04	-1.08	-1.16	-0.28	0.10	0.83	0.03	3.90	1.94	92.21
-0.46	0.11	-1.36	-0.43	-1.67	-0.47	-0.04	-1.83	-1.64	-0.22	-0.09	0.24	-0.25	-0.34	-3.17	-1.47	-32.29
-1.40	0.65	-1.14	-0.91	1.40	0.31	-0.43	0.24	0.30	-0.51	0.42	3.39	1.27	-0.76	-3.68	-1.78	-148.49
-0.74	2.37	0.65	-0.90	0.47	-0.88	-0.88	1.74	1.22	0.47	-1.93	-0.35	-0.49	1.63	0.56	-1.72	11.51
-0.41	0.35	-1.34	1.93	-1.60	-0.95	0.31	0.65	-0.21	1.54	-0.23	-0.33	0.28	-0.06	-3.10	-1.70	-114.79
1.85	-0.23	0.85	-0.13	1.27	0.37	-1.64	-0.24	-0.73	2.53	1.46	-0.40	0.67	1.61	3.54	1.08	171.13

Tabla 1: conjunto general de variables

Contesta a las siguientes preguntas:

1. (2 puntos) Representar los datos: analizar si hay posibles datos anómalos, tablas y gráficas a modo de resumen que se consideren relevantes. Divide el *dataset* en entrenamiento (0.8) + test (0.2). Interpreta y comenta los resultados obtenidos
2. (1.5 puntos) Realiza un ajuste de regresión lineal multivariante para predecir la variable y a partir del resto y comenta los resultados obtenidos
3. (1.5 puntos) Realiza una selección de variables con el método *stepwise* atendiendo al valor del VIF y comenta los resultados obtenidos
4. (1.5 puntos) Realiza una regresión con el método de LASSO y comenta los resultados obtenidos
5. (2 puntos) Compara los resultados obtenidos en los 3 ajustes atendiendo a los valores de R^2 , R^2 ajustado y el error cuadrático medio, además de a otros criterios a tu elección. Explica de forma razonada qué método crees que es mejor

Pregunta (1.5 puntos)

Explica en menos de 200 palabras el problema que puede acarrear tener datos anómalos en la muestra. Debes contestar a las preguntas ¿qué consecuencias tiene?, ¿cómo podemos detectarlo?, ¿cómo podemos trabajar si tenemos una muestra con datos anómalos?

DATOS PERSONALES		FIRMA
Nombre:	JUAN PABLO	DNI: 79953128
Apellidos:	POVEDA LOTERO	

[Firma manuscrita]

B O R R A D O R
PÁGINA NO VÁLIDA PARA RESPONDER

B O R R A D O R
PÁGINA NO VÁLIDA PARA RESPONDER