

TALLER REGRESIÓN LINEAL MULTIPLE

PRESENTADO POR:

JORGE ALBERTO INFANTE AVENDAÑO

JESUS DAVID SUAREZ PEÑA

PRESENTADO AL DOCENTE:

ALVARO AGUSTIN OÑATE BOWEN

UNIVERSIDAD POPULAR DEL CESAR

FACULTAD DE INGENIERÍA Y TECNOLOGÍAS

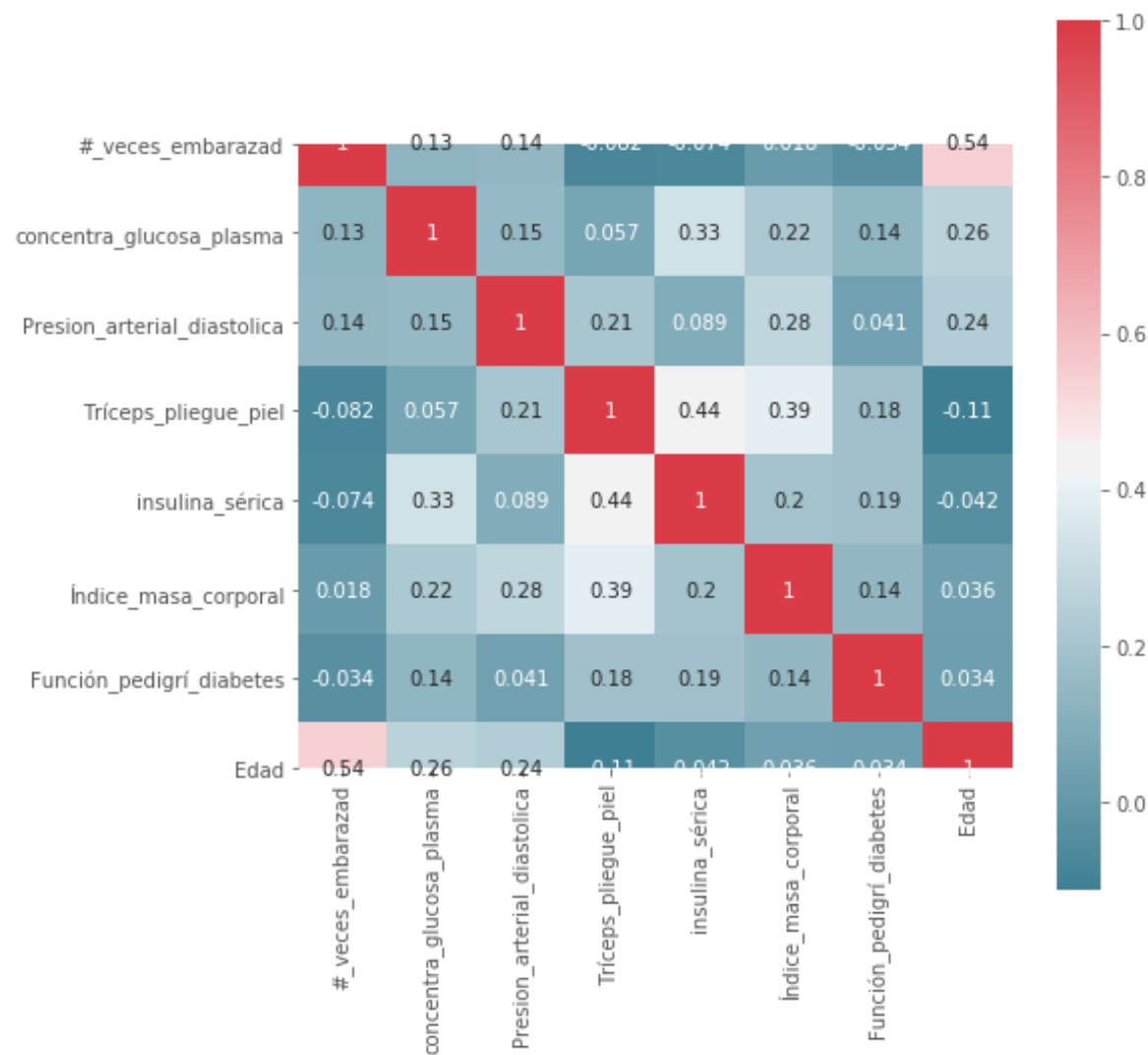
INGENIERIA DE SISTEMAS

VALLEDUPAR - CESAR

2020

1. Selecciona las variables de la base de datos y Calcular

Matriz de correlaciones



Covarianza

	0	1	2	3	4	5	6	7
0	11.3541	13.9471	9.21454	-4.39004	-28.5552	0.469774	-0.037426	21.5706
1	13.9471	1022.25	94.431	29.2392	1220.94	55.727	1.45487	99.0828
2	9.21454	94.431	374.647	64.0294	198.378	43.0047	0.264638	54.5235
3	-4.39004	29.2392	64.0294	254.473	802.98	49.3739	0.972136	-21.381
4	-28.5552	1220.94	198.378	802.98	13281.2	179.775	7.06668	-57.1433
5	0.469774	55.727	43.0047	49.3739	179.775	62.16	0.367405	3.36033
6	-0.037426	1.45487	0.264638	0.972136	7.06668	0.367405	0.109779	0.130772
7	21.5706	99.0828	54.5235	-21.381	-57.1433	3.36033	0.130772	138.303

Test Correlación Pearson

Index	#_veces_embaraza	entra_glucosa_pla	ion_arterial_diast	tríceps_pliegue_pi	insulina_sérica	índice_masa_corpor	función_pedigrí_diab	Edad
#_veces_emba...	1	0.129459	0.141282	-0.0816718	-0.0735346	0.0176831	-0.0335227	0.544341
concentra_glu...	0.129459	1	0.15259	0.0573279	0.331357	0.221071	0.137337	0.263514
Presion_arte...	0.141282	0.15259	1	0.207371	0.0889334	0.281805	0.0412649	0.239528
Tríceps_plie...	-0.0816718	0.0573279	0.207371	1	0.436783	0.392573	0.183928	-0.11397
insulina_sér...	-0.0735346	0.331357	0.0889334	0.436783	1	0.197859	0.185071	-0.042163
Índice_masa...	0.0176831	0.221071	0.281805	0.392573	0.197859	1	0.140647	0.0362419
Función_pedi...	-0.0335227	0.137337	0.0412649	0.183928	0.185071	0.140647	1	0.0335613
Edad	0.544341	0.263514	0.239528	-0.11397	-0.042163	0.0362419	0.0335613	1

2. Aplicar el Modelo de Regresión

Mostrar el Resultado del Modelo

En este caso para trabajar el datasets aplicando regresión lineal múltiple se escogió las siguientes variables las cuales son: Edad, insulina_sérica, Índice_masa_corporal

Coeficiente o Parámetros del Modelo

[2.91830579 -0.48408016]

Error cuadrático medio: 12712.31

Variancia: 0.04

T- Valor si es < 0,05 el Coeficiente es significativo

	coef	std err	t	P> t	[0.025	0.975]
const	2.5266	20.254	0.125	0.901	-37.233	42.286
x1	2.9183	0.518	5.637	0.000	1.902	3.935
x2	-0.4841	0.347	-1.395	0.164	-1.165	0.197

Niveles de Significancia

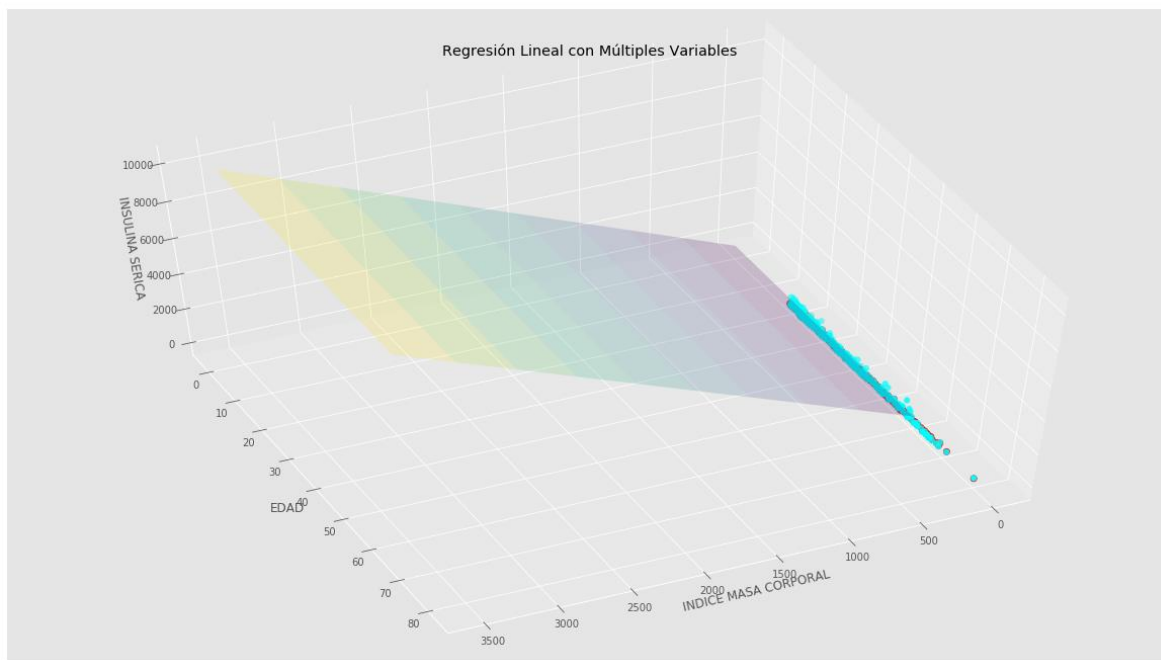
[9.00756298e-01 2.43603434e-08 1.63508661e-01]

R-squared

0.041585224796960696

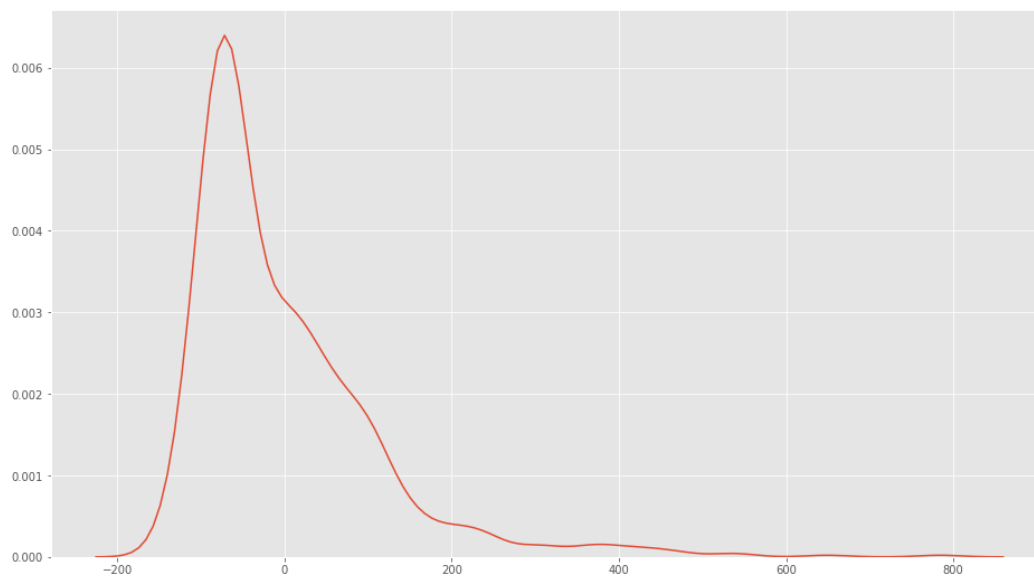
3. Realizar predicciones

Para realizar las predicciones en el modelo se escogió las variables Edad y Índice_masa_corporal la cual serán las que nos ayudara a predecir la dosis de insulina serial



5. Calcular los residuos del Modelo

Distribución de los residuos del modelo



Índice de masa corporal	Edad	Insulina
60	40	158
55	30	133
70	30	192