

Base de datos: neonatos
Variable de estudio: peso de los bebés (en gramos)
Variable categórica: Sexo del bebé
N de la Muestra: 178 chicos y 158 chicas
Resultados atípicos: 3 chicos y 1 chica.
N del estudio estadístico: 175 chicos y 157 chicas

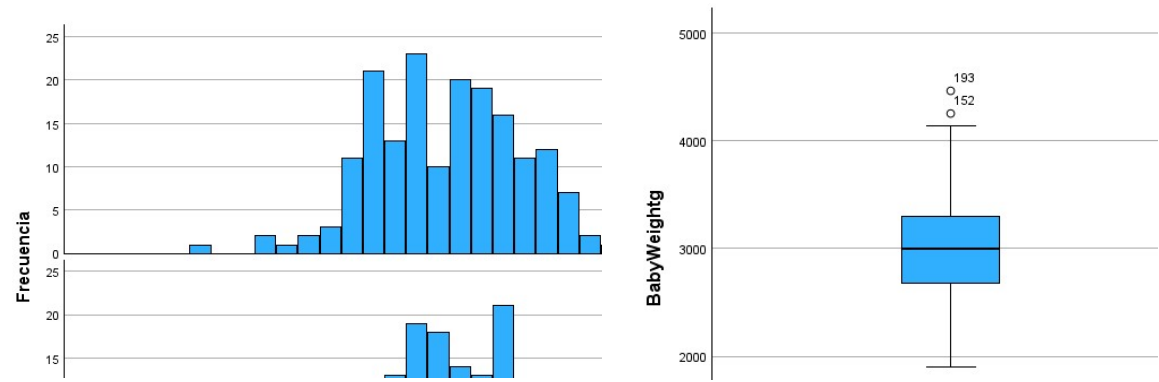


Figura 1: Izquierda, histograma de frecuencia absoluta del peso de los bebés separado por sexo. Derecha, Diagrama de cajas de las distribuciones del peso de los bebés separado por sexo. En las imágenes se muestran los datos sin depurar (incluye datos atípicos).

Descriptivos				Descriptivos			
Gender				Gender			
BabyWeightg	boy	Media		BabyWeightg	boy	Media	
		95% de intervalo de confianza para la media				95% de intervalo de confianza para la media	
		Límite inferior	Límite superior			Límite inferior	Límite superior
		Media recortada al 5%				Media recortada al 5%	
		Mediana				Mediana	
		Varianza				Varianza	
		Desv. estándar				Desv. estándar	
		Mínimo				Mínimo	
		Máximo				Máximo	
		Rango				Rango	
		Rango intercuartil				Rango intercuartil	
		Asimetría				Asimetría	
	Curtosis		Curtosis				
	girl	Media			girl	Media	
		95% de intervalo de confianza para la media				95% de intervalo de confianza para la media	
		Límite inferior	Límite superior			Límite inferior	Límite superior
		Media recortada al 5%				Media recortada al 5%	

Tabla 1: Resultados estadísticos para peso de los bebés separado por sexo. A la izquierda resultados para todos los datos de la muestra, a la derecha extraemos los datos atípicos.

Vemos que las conclusiones cambian cuando extraemos los datos atípicos. Interpretamos los datos depurados:

- La media de los chicos (2988.15g) es mayor que la de las chicas (2984.91g). La distribución de los chicos presenta una desviación estándar de 418.47g mientras que la de las chicas es de 409.58g. Por lo que **las chicas presentan una media más representativa** (cv=0.137) que los chicos con (cv=0.140) siendo la distribución de los chicos la más dispersa, tanto de relativa a su media como absoluta.
- La distribución de las chicas es **un poco asimétrica hacia la izquierda** (-0.327) y **un poco platicurtica** (-0.215).
- La distribución de los chicos es **un poco asimétrica hacia la derecha** (0.051) y **un poco platicurtica** (-0.364).
- Ambas distribuciones presentan asimetrías y curtosis dentro de los parámetros normales ($-2 < g_1, g_2 < 2$) por lo que podemos considerarlas como gaussianas.

Base de datos: neonatos
Variable de estudio: peso de los bebés (en gramos)
Variable de estudio: peso de las madres (en kilogramos)
Variable categórica: Sexo del bebé
N de la Muestra: 178 chicos y 158 chicas

Regresión lineal de peso del bebé sobre peso de la madre

y=ax+b	Sin separación de genero	boy	girl
a=	10.888	9.070	13.316
b=	2357.608	2459.018	2223.470
r^2=	0.111	0.080	0.160

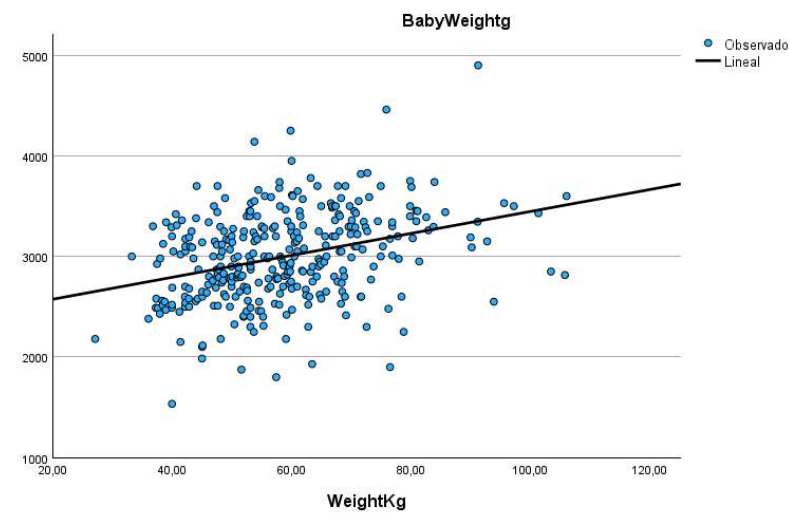


Figura 2: Regresión lineal de peso del bebé sobre peso de la madre sin separación de género.

- Parece que las variables no están fuertemente relacionadas ya que r^2 es un valor pequeño, lo que indica una gran dispersión en las varianzas de los datos.
- Se podría intuir una relación creciente suave, a más peso de la madre más peso del bebé, sin embargo, las predicciones conllevarían un error del orden de $1-r^2=84\%$ que es muy grande.
- Parece haber diferencias entre sexos. Las chicas muestran más dependencia con el peso de la madre que los chicos ya que tienen una r^2 mayor.
- Las chicas también muestran una pendiente mayor, lo que indicaría que su peso al nacer está más relacionado con el peso de la madre que para los chicos.
- El término independiente es mayor para los chicos, ¿qué interpretación le podríamos dar?

Base de datos: neonatos
Variable de estudio: peso de las madres ‘WeightMeasures’ (en kilogramos), en distintos momentos del embarazo (V4,V5, V6, V7, V8))
Variable categórica: Sexo del bebe
N de la Muestra: 178 chicos y 158 chicas

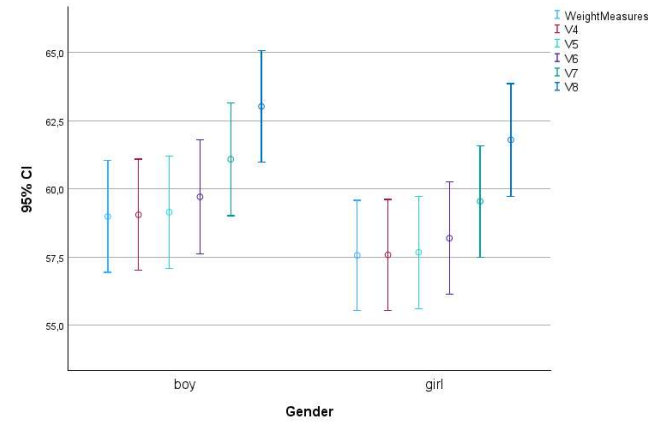


Figura 3: Intervalos de confianza al 95% para la comparación de la media de los pesos de las madres en distintos momentos del embarazo (WeightMeasures , V4,V5, V6, V7, V8) y separando por género del bebe

1. Comparación entre la misma población: todas las madres

Prueba de muestras emparejadas									
Diferencias emparejadas					Significación				
		Media	Desv. estándar	Media de error estándar	95% de intervalo de confianza de la diferencia		t	gl	P de un factor
Par 1	WeightMeasures - V8	-4,13277	3,26910	,17834	Inferior	Superior	-23,173	335	<,001
					-4,48358	-3,78195			<,001

Vemos **diferencia significativa** al 95% entre las dos medidas de peso WM, V8 ya que el **intervalo de confianza para la diferencia NO cruza al cero**: son medias significativamente distintas.

Prueba de muestras emparejadas									
Diferencias emparejadas					Significación				
		Media	Desv. estándar	Media de error estándar	95% de intervalo de confianza de la diferencia		t	gl	P de un factor
Par 1	V4 - V5	-,09435	1,24473	,06791	Inferior	Superior	-1,389	335	,083
					-,22792	,03923			,166

Vemos **igualdad significativa** al 95% entre las dos medidas de peso V4 y V5 ya que el **intervalo de confianza para la diferencia SI cruza al cero**: son medias significativamente iguales.

2. Comparación entre distintas poblaciones: madres de niños vs madres de niñas

Prueba de muestras independientes										
Prueba de Levene de igualdad de varianzas				prueba t para la igualdad de medias						
		F	Sig.	t	gl	Significación	P de un factor	P de dos factores	Diferencia de medias	Diferencia de error estándar
V4	Se asumen varianzas iguales	,002	,968	-1,004	334	,158	,316	,316	-1,47025	1,46457
	No se asumen varianzas iguales			-1,008	332,910	,157	,314	,314	-1,47025	1,45914
				Inferior	Superior	70% de intervalo de confianza de la diferencia				
				-2,99054	,05004					
				-2,98491	,04440					

Vemos **igualdad significativa** al 70% entre la medida de peso V4 separando para madres de chicos y chicas, ya que el **intervalo de confianza cruza al cero**: son medias **significativamente iguales** al 70%.

Prueba de muestras independientes										
Prueba de Levene de igualdad de varianzas				prueba t para la igualdad de medias						
		F	Sig.	t	gl	Significación	P de un factor	P de dos factores	Diferencia de medias	Diferencia de error estándar
V4	Se asumen varianzas iguales	,002	,968	-1,004	334	,158	,316	,316	-1,47025	1,46457
	No se asumen varianzas iguales			-1,008	332,910	,157	,314	,314	-1,47025	1,45914
				Inferior	Superior	50% de intervalo de confianza de la diferencia				
				-2,45917	-,48134					
				-2,45550	-,48500					

Vemos **diferencia significativa** al 50% entre la medida de peso V4 separando para madres de chicos y chicas, ya que el **intervalo de confianza no cruza al cero**: son medias **significativamente distintas** al 50%.