

Estimación de la altura a partir de restos óseos

Contexto

El descubrimiento de restos óseos plantea numerosas preguntas que son mejor respondidas por aquellos capacitados en antropología forense.

Los antropólogos forenses no solo son capaces de determinar en el sitio si los restos son humanos, sino que también emplean varios métodos para determinar el sexo, la edad al momento de la muerte, la raza y la altura del fallecido.

La estatura se considera exclusiva de la especie humana, ya que los animales no asumen una postura erguida habitual fisiológica, y sus dimensiones dependen de varios segmentos del esqueleto, como el cefálico, raquídeo (altura de la columna), pelviano y de los miembros inferiores, cada uno contribuye a la estatura del individuo en consideración a la edad, sexo, raza, condiciones socioeconómicas y psicosociales, y finalmente las tendencias históricas (Rodríguez, 1994). Para Krishan et al. (2010), la estimación de la estatura es un parámetro importante en la identificación de los restos óseos durante los exámenes forenses.

El tema de la estatura es complejo y variable, en el que intervienen muchos factores: genéticos (herencia multifactorial), sexo (es un parámetro biológico de dimorfismo sexual puesto que en todos los grupos humanos, la estatura de la mujer suele ser inferior a la del hombre), edad, condiciones ambientales como la alimentación, la situación socioeconómica, el medio geográfico, etc. Por otro lado, está influido por enfermedades óseas (patologías), e incluso es cambiante dentro de un mismo individuo en diferentes momentos del día.

La relación entre la estatura de un individuo y la longitud de sus huesos largos permite calcular la talla a partir de éstos.

Diversos autores han señalado que las estimaciones de la estatura que se realizan, teniendo como base la medida de los huesos largos de las extremidades inferiores (fémur, tibia, peroné), son más precisas que las estimaciones realizadas con huesos largos de las superiores (húmero, cúbito, radio). Concretamente, Mendonça (1998) señala que en la población portuguesa, la estimación de la estatura con la longitud del fémur es más correcta que con la del húmero. Bhavna y Surinder Nath (2007), en su estudio con población india, señalan que la tibia es el hueso que proporciona mejores estimaciones con una menor desviación estándar y un coeficiente de correlación r mayor que el resto de los huesos estudiados. Meadows y Jantz (1999), con población americana, hallaron una relación más estrecha entre la estatura y

longitud de tibia o peroné, que entre estatura y longitud de fémur. Agnihotri, et al. (2009) igualmente concluyeron en su estudio con población indo mauritana, que la tibia tiene una correlación más fuerte con la estatura que el cúbito.

Los intentos científicos por estimar la estatura a partir de los restos esqueléticos cuentan con una larga historia y en general se han seguido dos enfoques principales:

El método matemático consiste en medir la longitud de una o varias combinaciones de huesos largos y usar esto para predecir la estatura y en segundo lugar está el método anatómico en el cual se usa el esqueleto completo donde se suman todas las alturas de los segmentos óseos que contribuyen a la estatura (İsçan & Steyn, 2013): los segmentos medidos son el cráneo, la vértebras, el fémur, la tibia, el astrágalo y el calcáneo (Dayala, et al. 2008; Chibba & Bidmos, 2007) y a esto se le suma el valor que compensa los tejidos blandos para así estimar la altura. Se ha planteado que la inclusión de la altura de todos los huesos que contribuyen a la altura, elimina las imprecisiones introducidas por diferentes proporciones corporales, que es una de las principales fuentes de error del método matemático (Brits, et al. 2017).

En 1894, Dwight introdujo el método anatómico, donde estimó la altura esquelética total midiendo todos los elementos que constituyen la altura, este método fue reintroducido por Fully y hoy se le conoce como el método completo o anatómico, al estimar la estatura mediante este método se ha considerado que se puede usar en cualquier tipo de población con los factores de corrección de los tejidos blandos que son comunes a todas las poblaciones (Dayala, et al. 2008), sin embargo se le asocia un nivel de imprecisión asociado con las descripciones de cómo medir los huesos, así pues en una revisión del método Raxter et al. (2006) abordaron estas cuestiones y proporcionaron explicaciones y descripciones de cómo medir los segmentos individuales y también proporcionaron nuevos factores de corrección de los tejidos blandos ya que los propuestos por Fully subestimaban la estatura (Baines, et al. 2016).

Esta versión revisada del método anatómico ha sido utilizada para desarrollar o evaluar fórmulas de regresión en investigaciones en las que no se cuenta con una documentación certera de la estatura en vida de los individuos pertenecientes a las muestras (Brits, et al. 2017), sin embargo la mayor limitación para utilizar este método en otro tipo de contextos como el arqueológico o el forense, es la incidencia de elementos óseos faltantes además se haya una serie de factores que contribuyen en la preservación de los restos humanos (Auerbach, 2011; Jeong & Jantz, 2016). Dada esta situación es que se recurre con mayor frecuencia al uso de los métodos matemáticos.

Entre los pioneros en el estudio de la estimación de la estatura basada en los restos óseos se encuentra a Sue quien en el año 1755 presentó algunas conclusiones sobre la proporcionalidad del esqueleto humano según el estudio que realizó en 14 cadáveres de diferentes edades; más tarde en 1821 Orfila hizo un estudio midiendo diez esqueletos y 51 cadáveres de París de individuos franceses y publicó las que se consideran las primeras tablas de determinación de la estatura a partir de los huesos largos (Belmonte, 2013).

En 1888 Rollet publicó las primeras tablas sobre la estatura utilizando el húmero, el radio, el cúbito, el fémur, la tibia y el peroné en una muestra de 50 hombres y 50 mujeres de cadáveres franceses, midiendo los huesos en estado fresco y luego en estado seco, para este momento habían perdido 2 milímetros de longitud. En 1892 y 1893, Manouvrier revaluó los datos de Rollet y excluyó a todos los sujetos mayores de 60 años ya que Manouvrier planteaba que a partir de esta edad, estos perdían unos 3 centímetros de estatura (İsçan & Steyn 2013). En sus cálculos también tomó en consideración las diferencias en las proporciones de las extremidades y la talla en función del cuerpo así como la influencia de ciertos procesos en la longitud del mismo, además fue el primero en afirmar que la estatura es una función de la longitud de los huesos largos (Hauser, et al. 2005).

Sin embargo estos autores calculaban estas medidas como parte de sus investigaciones sobre las proporciones corporales en poblaciones humanas pero ninguno las usó con fines de calcular la estatura (Feldesman, et al. 1990). Ya para 1899 utilizando los datos de Rollet, Pearson realizó el primer intento de predecir la estatura a partir de las mediciones de los huesos largos (Stevenson, 1929) y desarrolló fórmulas de regresión, basadas en los huesos del lado derecho y también estableció reglas básicas para su reconstrucción, que aún siguen siendo vigentes. Desde entonces se han publicado varias ecuaciones que usan una amplia variedad de huesos para predecir la estatura en vida (İsçan & Steyn 2013). Este método es un modelo matemático predominante también conocido como calibración inversa en donde se usa la regresión estadística (Moore, M. K. & Ross, A. H. 2012). En 1937, Breitinger propuso un método adicional para calcular la estatura a partir de las mediciones de los huesos largos, realizó sus estudios entre los habitantes de Europa Central utilizando medidas realizadas sobre sujetos vivos, empleando radiografías para este propósito (Hauser, et al. 2005).

En general, dada su indudable utilidad práctica, tradicionalmente se han utilizado las tablas de Manouvrier (1893), elaboradas mediante medidas tomadas de individuos franceses, y las tablas y fórmulas regresivas de Trotter y Gleser (1952, 1958, 1970) con población americana.

En la actualidad, en la antropología forense americana se utiliza el programa informático FORDISC 3.0 (Jantz y Ousley, 1993-2005), que se nutre de la Forensic Anthropology Data Bank (que contiene muestras representativas de los diferentes

grupos étnicos que conviven en Estados Unidos) para proporcionar estimaciones de sexo, origen poblacional y fórmulas regresivas para estimar la estatura.

Caso de Estudio

El siguiente caso está basado en un estudio publicado por la University of Southern Denmark, titulado 'Height Estimation from Skeletal Remains', cuyo autor es Mette Wod. El mismo forma parte de un proyecto de biomedicina sobre la determinación de la altura a partir de información de medidas de restos óseos antiguos encontrados en el pueblo Ribe de Dinamarca

Los esqueletos utilizados como material en este análisis fueron excavados en un cementerio medieval en Ribe en Jutlandia, Dinamarca en 1993.

Un total de 429 fosas fueron excavadas y la altura en la tumba pudo ser medida en 184 individuos. De las 184 personas, 114 personas (67 hombres y 47 mujeres) fueron adultos y apropiados para este análisis. A su vez, la edad y el sexo fueron determinados para todos los individuos mediante métodos osteológicos.

El propósito de este trabajo es analizar los datos de los huesos extraídos y examinar la asociación entre la altura adulta y la longitud del fémur mediante análisis de regresión.

Además, se investiga si los períodos de debilidad en la infancia, en este trabajo de la edad 1-6 inclusive, que se manifiestan como hipoplasias, tienen un impacto en la altura y las proporciones corporales de los adultos.

Dataset

El conjunto de datos se encuentra en un archivo de Microsoft Access, disponible en <http://mirror.aidoh.dk/mettewod/wod.zip>


El archivo se cargó con Microsoft Excel, se seleccionó la tabla con los datos y luego se convirtió a un CSV.

Los datos que se incluyen son de adultos cuya altura en la tumba es conocida. Un adulto se define como un individuo donde la sincondrosis esfenoccipital (SOS) y todas las epífisis de los huesos largos se han fusionado.

Si no hay datos disponibles para un atributo, el campo se deja vacío.

Row No.	ID	Location	Site_Number	Age_Minum...	Age_Maxim...	Sex	Grave_Num...	Canine num...	Canine_larg...	Canine_2nd...	Incisor_num...	Inciso_larg...	Incisor_2nd...	Height_in_g...	Abnormaliti...	Femur...
1	1	Ribe	ASR1015	20	24	Male	G40	2	3	5	5	3	3.500	173.500	?	49.600
2	4	Ribe	ASR1015	35	45	Male	G312	3	4	4.500	3	2	3	170	?	48.400
3	5	Ribe	ASR1015	50	60	Male	G229	0	?	?	0	?	?	171.500	Two vertebra...	50.800
4	6	Ribe	ASR1015	30	40	Male	G257	?	?	?	?	?	?	165	?	45.400
5	7	Ribe	ASR1015	45	55	Male	G74	?	?	?	?	?	?	165	?	47.500
6	8	Ribe	ASR1015	36	40	Male	G61	?	?	?	?	?	?	164	?	47
7	9	Ribe	ASR1015	50	60	Male	G377	3	3.500	5	?	?	?	166	?	?
8	10	Ribe	ASR1015	55	65	Male	G267	?	?	?	?	?	?	185	?	52
9	11	Ribe	ASR1015	27	30	Male	G39	1	4.500	?	0	?	?	182	?	51.400
10	12	Ribe	ASR1015	50	60	Female	G104	?	?	?	?	?	?	150	?	?
11	13	Ribe	ASR1015	35	44	Female	G159	2	4	4.300	4	2.500	3.200	158.500	?	?
12	14	Ribe	ASR1015	20	30	Male	G851	?	?	?	?	?	?	164	?	?
13	15	Ribe	ASR1015	35	50	Male	G89	?	?	?	?	?	?	154	?	44.900
14	16	Ribe	ASR1015	45	55	Female	G248	?	?	?	?	?	?	164	?	?
15	17	Ribe	ASR1015	35	45	Male	G213	?	?	?	?	?	?	169	growing, but ...	45.600
16	18	Ribe	ASR1015	40	60	Male	G314	?	?	?	?	?	?	170	?	47.300
17	19	Ribe	ASR1015	35	45	Male	G435	?	?	?	?	?	?	167	?	?
18	20	Ribe	ASR1015	30	40	Male	G317	2	3	4.500	3	2.500	3	166	?	45.600
19	21	Ribe	ASR1015	36	44	Male	G387	?	?	?	?	?	?	160	?	?
20	22	Ribe	ASR1015	35	45	Female	G38	?	?	?	?	?	?	154	?	43.600
21	23	Ribe	ASR1015	25	34	Male	G368	4	3	3.500	5	3.500	3	169	?	?
22	24	Ribe	ASR1015	36	44	Male	G282	2	5	4.500	?	?	?	157.500	Some vertebr...	44.100
23	25	Ribe	ASR1015	30	35	Male	G206	4	3.500	3	4	2	1.500	156	?	43.400
24	26	Ribe	ASR1015	25	35	Male	G432	?	?	?	0	?	?	148	?	41.600
25	28	Ribe	ASR1015	40	50	Male	G166	5	2.500	3.500	?	?	?	170	?	48.300

Descripción de atributos



Retrieve Dataindsamling2.output (output)

Meta data: Data table

- Source: //Local Repository/data/Dataindsamling2

Number of examples = 117

23 attributes:

Note: Some of the nominal values in this set were discarded due to performance reasons. You can change this behaviour in the preferences (rapidminer.general.md_nominal_values_limit).

Generated by: Retrieve Dataindsamling2.output

Data: IOTable: 117 examples, 23 regular attributes, no special attributes

Role	Name	Type	Range	Missings
	ID	# integer	=[1 – 120]	= 0
	Location	nominal	=[G216, Ribe]	= 0
	Site_Number	nominal	=[ASR1015]	= 0

La imagen muestra que el dataset tiene varios valores faltantes que deberán ser tratados apropiadamente.

Hay varios registros en los que falta el dato del fémur izquierdo o el derecho. En estos casos se puede utilizar la medida del hueso existente ya que se ha determinado mediante un T-test que no hay una diferencia significativa entre ambos huesos y cualquiera de ellos es válido para utilizar en el estudio. En los casos en que faltan ambos valores, el registro se descarta.

A su vez hay dos columnas que contienen notas o comentarios acerca de las características de los restos encontrados, y si son válidos para utilizar en el estudio. En el caso que no lo sean se eliminan del dataset.

Selección de atributos

Se eligen los atributos que sean más relevantes a la hora de analizar la posible relación lineal o predecir la altura de la persona.

La lista de atributos que permanecen es la siguiente:

Age_Minimum, Age_Maximum, Sex, Height_in_grave, Femurleft, Femurright, Hyperplasia

```
bones = pd.read_csv("height_estimation.csv")
```

```
bones_subset = bones[['Age_Minimum', 'Age_Maximum', 'Sex', 'Height_in_grave',  
'Femurleft', 'Femurright', 'Hyperplasia']]
```

```
male_subset = bones_subset.loc[bones_subset["Sex"] == "Male"]
```

```
female_subset = bones_subset.loc[bones_subset["Sex"] == "Female"]
```

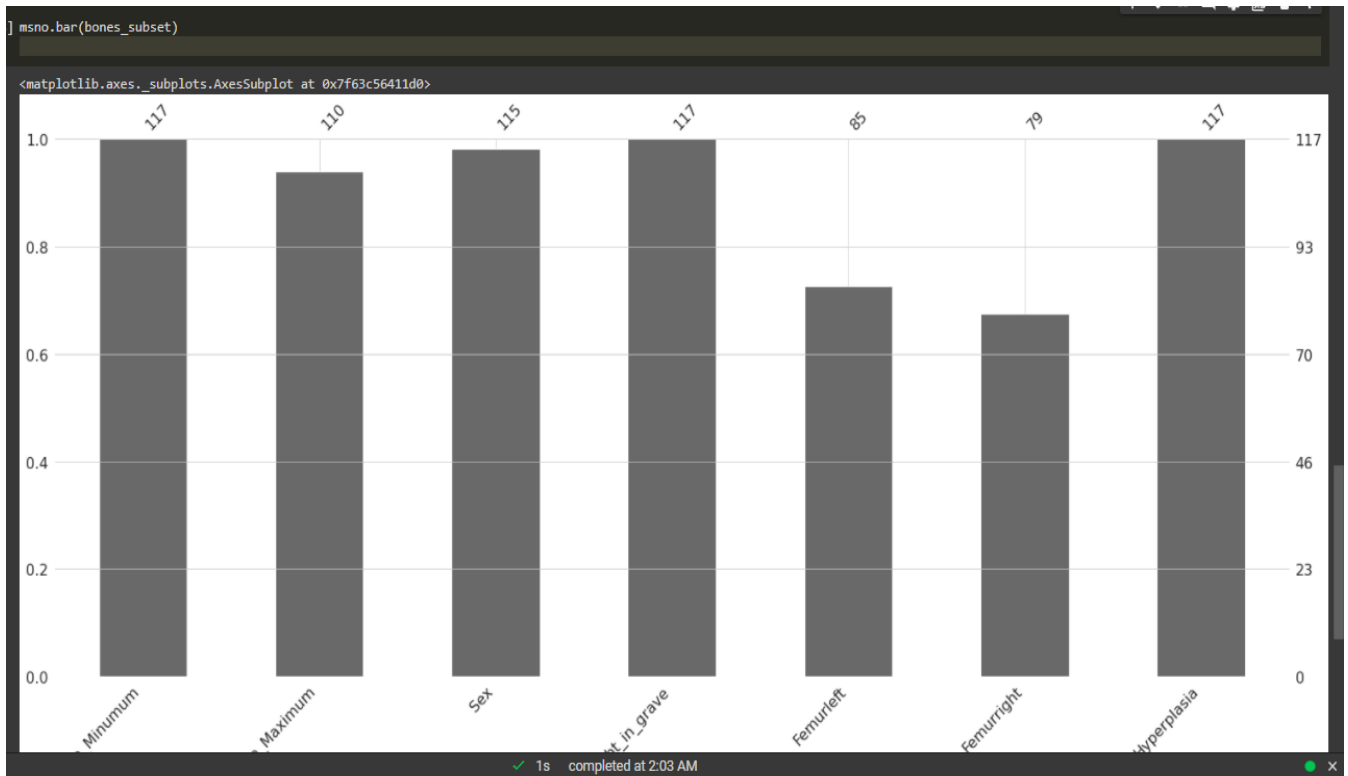
Datos faltantes

Como se ve en la siguiente gráfica, la mayor parte de los datos faltantes se produce en las medidas de los huesos. Como se mencionó anteriormente esto se soluciona asignando el valor del hueso existente. Si faltan ambas medidas se elimina.

Hay dos ejemplos que no tienen sexo. En las notas que se ingresaron sobre estos ejemplos se indica que no deberían calificar como registros válidos por lo cual se eliminan del dataset.

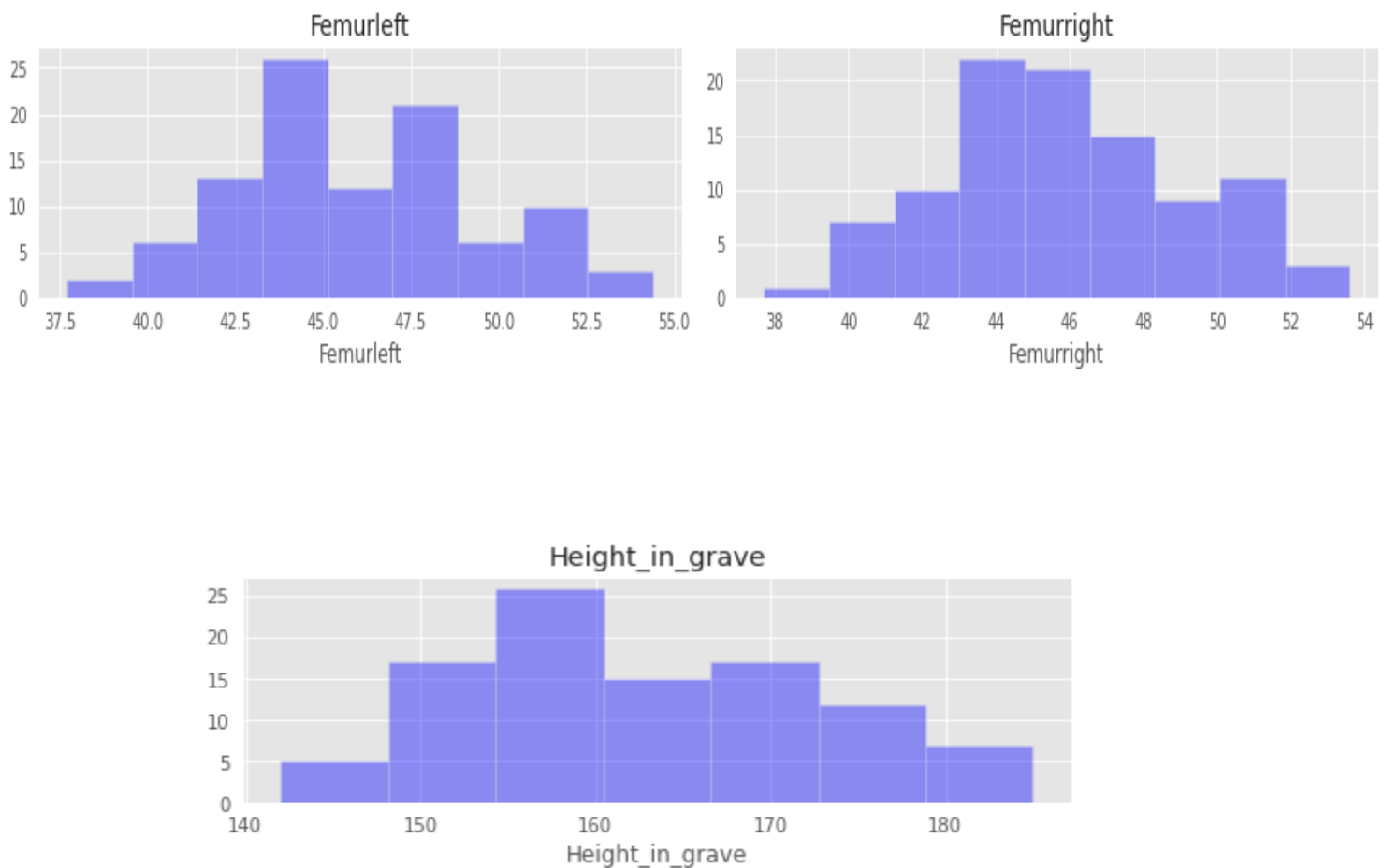
```
msno.bar(bones_subset)
```

-



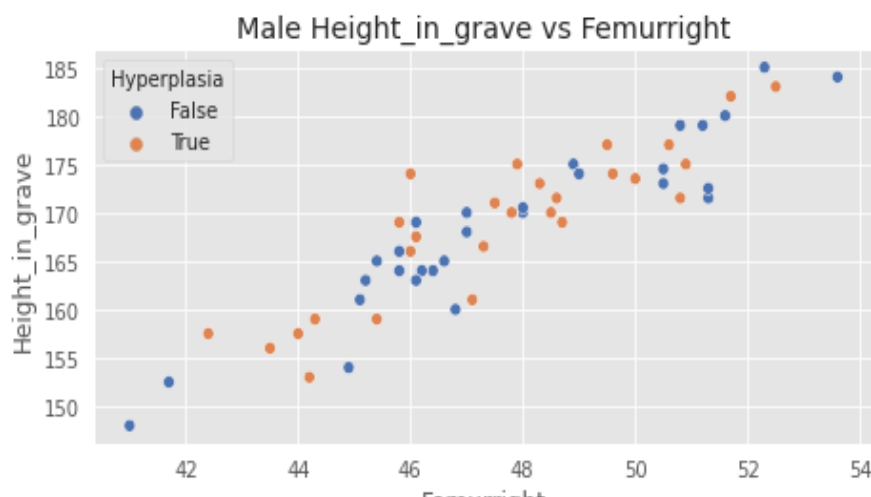
Visualización de datos

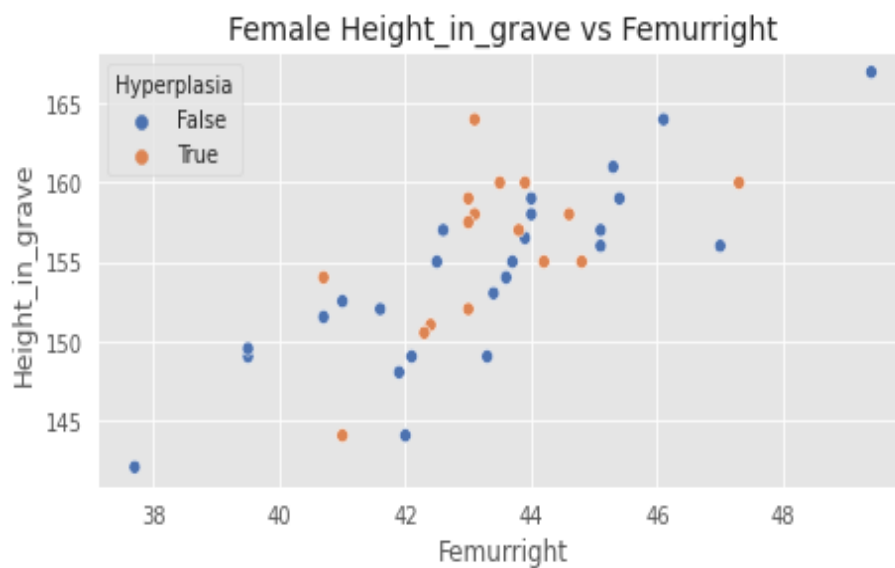
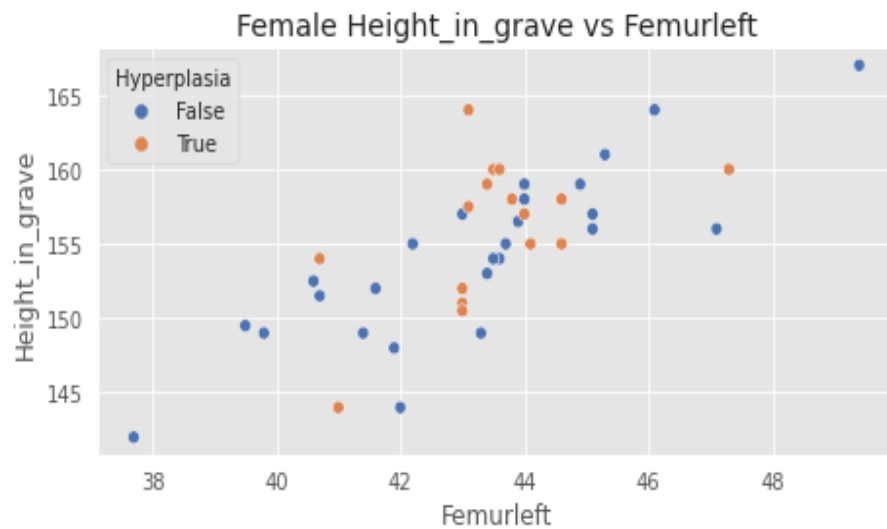
Mediante librerías de Python se visualizan una serie de gráficos para entender mejor los datos



Los atributos Femurleft, Femurright y Height_in_Grave tienen una distribución casi normal. Las medidas del fémur derecho parecen tener una mejor distribución que las del lado izquierdo.

Se realizaron varios diagramas de dispersión, con la longitud del fémur en función de la altura, para buscar tendencias, valores atípicos, fallas de medición, etc y para obtener una descripción general de la asociación entre estos.





Los diagramas de dispersión no mostraron signos de outliers, pero sí una clara tendencia lineal positiva. Se grafica también si los individuos tienen o no hiperplasia, para determinar si se producen variaciones en los datos de la muestra, Test estadísticos lo confirmarán, pero no existen diferencias en las estaturas de individuos con o sin hiperplasia.

Como se mencionó más arriba, no siempre es posible que ambos fémures estén disponibles para el análisis. Para poder utilizar a un individuo con sólo un fémur disponible, debe comprobarse si los dos fémures difieren en longitud. Para hacer

esto se realizó una prueba t de muestras entre los fémures izquierdo y derecho. Se realizó la prueba t con un nivel de significancia de 0.05, y no mostró diferencias. Esto significa que no hay una diferencia significativa entre los dos fémures. Es decir, si bien es preferible utilizar la media para completar los datos faltantes, es estadísticamente válido usar solo el fémur disponible.

```
left_men = bones.loc[bones["Sex"] == "Male", "Femurleft"]
right_men = bones.loc[bones["Sex"] == "Male", "Femurright"]
left_women = bones.loc[bones["Sex"] == "Female", "Femurleft"]
right_women = bones.loc[bones["Sex"] == "Female", "Femurright"]

def independent_ttest(data1, data2, alpha):
    # calculate means
    mean1, mean2 = mean(data1), mean(data2)
    # calculate standard errors
    se1, se2 = sem(data1, nan_policy='omit'), sem(data2, nan_policy='omit')
    # standard error on the difference between the samples
    sed = sqrt(se1**2.0 + se2**2.0)
    # calculate the t statistic
    t_stat = (mean1 - mean2) / sed
    # degrees of freedom
    df = len(data1) + len(data2) - 2
    # calculate the critical value
    cv = t.ppf(1.0 - alpha, df)
    # calculate the p-value
    p = (1.0 - t.cdf(abs(t_stat), df)) * 2.0
    # return everything
    return t_stat, df, cv, p

alpha = 0.05
t_stat, df, cv, p = independent_ttest(left_men, right_men, alpha)
print('t=%.3f, df=%d, cv=%.3f, p=%.3f' % (t_stat, df, cv, p))
# interpret via p-value
if p > alpha:
    print('Males: Accept null hypothesis that the means are equal.')
else:
    print('Males: Reject the null hypothesis that the means are equal.')
```

```
t_stat, df, cv, p = independent_ttest(left_women, right_women, alpha)
print('t=%.3f, df=%d, cv=%.3f, p=%.3f' % (t_stat, df, cv, p))

if p > alpha:
    print('Females: Accept null hypothesis that the means are equal.')
else:
    print('Females: Reject the null hypothesis that the means are equal.')

t=0.168, df=112, cv=1.659, p=0.867
Males: Accept null hypothesis that the means are equal.
t=0.045, df=82, cv=1.664, p=0.964
Females: Accept null hypothesis that the means are equal.
```

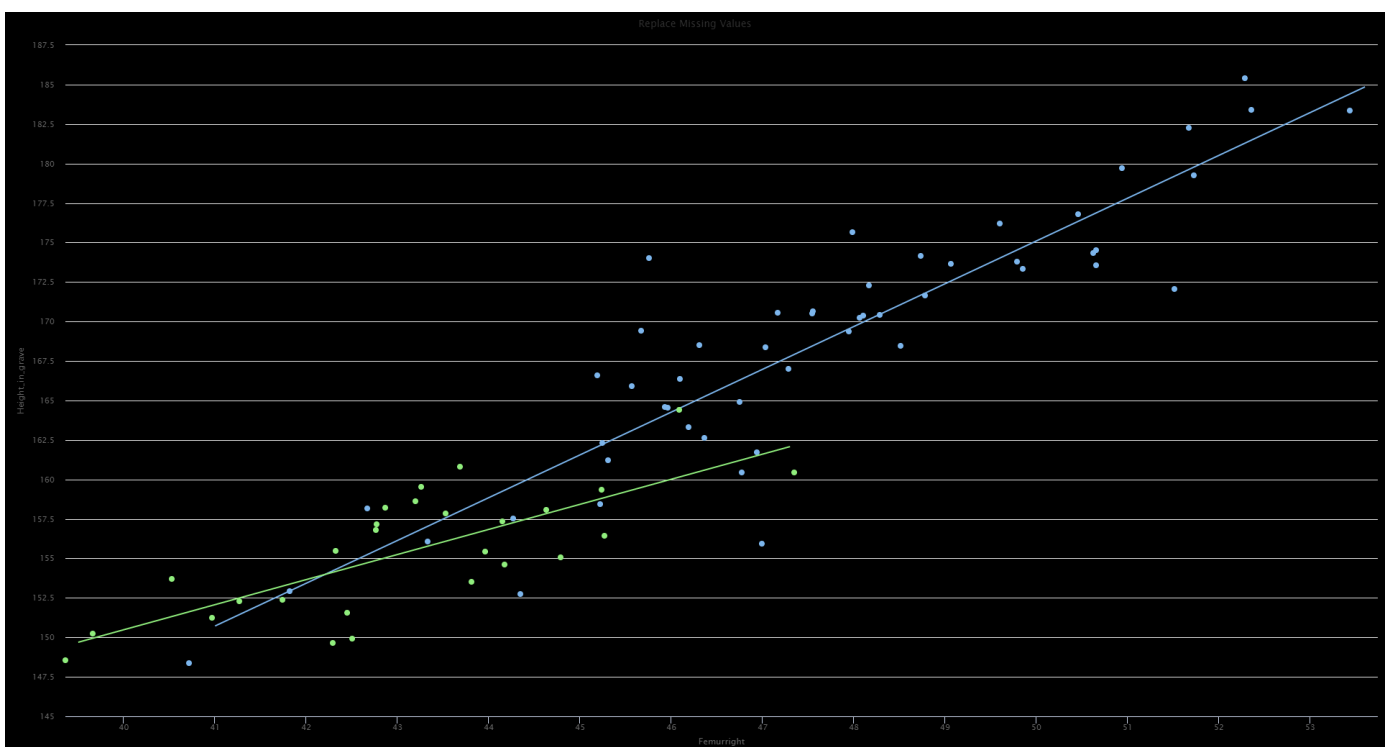
Análisis de regresión

Tanto la bibliografía como la visualización previa de datos sugieren una relación lineal entre el largo del fémur derecho o izquierdo con la altura de la persona.

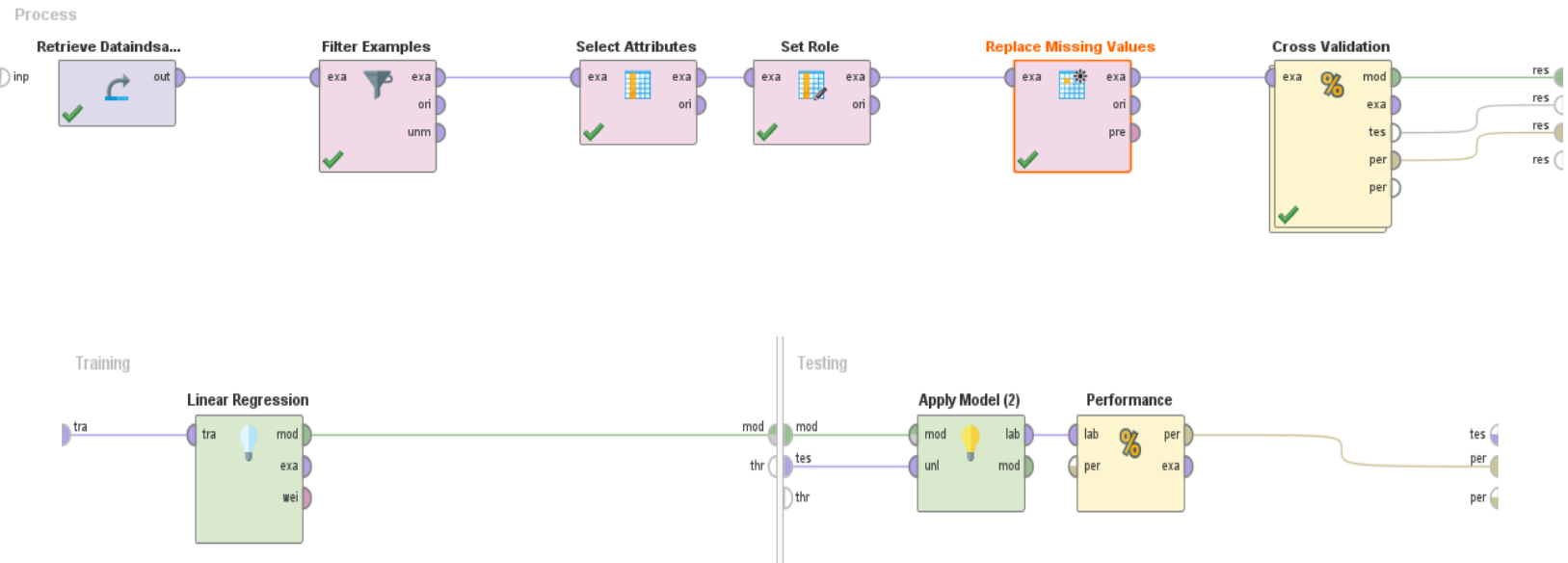
Para verificarlo, cargamos el dataset en RapidMiner y realizamos las operaciones descritas anteriormente para preparar los datos y cargar un modelo de regresión lineal.

El análisis se hace por separado para hombres y mujeres, con y sin hiperplasia, cambiando el filtro correspondiente.

Graficando los datos e interpolando una regresión lineal vemos que se aproximan muy bien a ese modelo. :



Luego aplicamos el modelo de regresión lineal simple a los datos con RapidMiner:






El resultado de la regresión es el siguiente:

PerformanceVector (Performance)

LinearRegression (Linear Regression)

Attribute	Coefficient	Std. Error
femur_length	2.662	0.277
(Intercept)	41.544	13.229

 PerformanceVector (Performance)		 LinearRegression (Linear Regression)
-------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------

Attribute	Coefficient	Std. Error
femur_length	1.901	0.249
(Intercept)	71.958	10.765

Las ecuaciones que obtiene el modelo son las siguientes:

Hombres: $\text{Height} = 2.662 \times \text{femur_length} + 41.544$

Mujeres: $\text{Height} = 1.901 \times \text{femur_length} + 71.958$

Estos valores son bastante similares a los obtenidos por el estudio, los cuales son:

The linear regressions are as follows:

$$\text{Height}_{\text{men}} = 2.610 \cdot \text{femur} + 44.201$$

$$\text{Height}_{\text{women}} = 2.019 \cdot \text{femur} + 67.579$$