

**ANÁLISIS DE DATOS
LABORATORIO 1: ANÁLISIS ESTADÍSTICO**

Autores:

Nicolás Mariángel Toledo

Juan Pablo Rojas Rojas

Profesor:

Max Chacón Pacheco

Ayudante:

Ignacio Ibáñez Aliaga

TABLA DE CONTENIDOS

ÍNDICE DE FIGURAS.....	iv
ÍNDICE DE TABLAS.....	v
CAPÍTULO 1. INTRODUCCIÓN.....	7
1.1 MOTIVACIÓN	7
1.2 OBJETIVOS	7
1.3 ORGANIZACIÓN DEL DOCUMENTO	7
CAPÍTULO 2. DESCRIPCIÓN DEL PROBLEMA.....	9
2.1 DESCRIPCIÓN DE LA BASE DE DATOS	10
2.1.1 Descripción de clases y variables	10
CAPÍTULO 3. ANÁLISIS ESTADÍSTICO E INFERENCIAL.....	13
3.1 Análisis de variables booleanas	14
3.1.1 Género	14
3.1.2 Embarazo	14
3.1.3 Litio	15
3.1.4 Operado de la tiroides	16
3.1.5 Psych	16
3.1.6 Bocio (Goitre)	17
3.1.7 Tumor	17
3.2 Análisis de variables continuas	19
3.2.1 Medidas de tendencia central y dispersión	19
3.2.2 Relación entre variables	24
CAPÍTULO 4. CONCLUSIONES.....	25
CAPÍTULO 5. BIBLIOGRAFÍA.....	27

ÍNDICE DE FIGURAS

2.1	Condición de tiroides según relación T4-TSH [1].	9
3.1	Cantidad de pacientes de acuerdo a su diagnóstico.	13
3.2	Resultados del diagnóstico de acuerdo al sexo de los pacientes.	14
3.3	Resultados del diagnóstico de acuerdo a si la paciente se encuentra embarazada.	15
3.4	Resultados del diagnóstico de acuerdo a si el paciente ha sido medicado con litio.	15
3.5	Resultados del diagnóstico para sólo los pacientes que han sido medicados con litio.	16
3.6	Resultados del diagnóstico dado que el paciente ha pasado por cirugías que implique la tiroides.	16
3.7	Resultados del diagnóstico dado que el paciente se encuentre en algún tratamiento psiquiátrico.	17
3.8	Resultados del diagnóstico dado que el paciente padezca de bocio.	17
3.9	Resultados del diagnóstico dado que el paciente posea algún tipo de tumor en la glándula tiroides.	18
3.10	Distribución y medidas de tendencia central de la edad.	19
3.11	Distribución y medidas de tendencia central de TSH.	20
3.12	Acercamiento sin cola a la derecha de la distribución y medidas de tendencia central de TSH.	20
3.13	Distribución y medidas de tendencia central de T3.	21
3.14	Distribución y medidas de tendencia central de TT4.	21
3.15	Distribución y medidas de tendencia central de T4U.	22
3.16	Distribución y medidas de tendencia central de FTI.	22

ÍNDICE DE TABLAS

Tabla 3.1: Medidas de tendencia central y dispersión de la muestra.	19
Tabla 3.2: Medidas de tendencia central y dispersión de hombres de la muestra.	23
Tabla 3.3: Medidas de tendencia central y dispersión de mujeres no embarazadas de la muestra. . .	23
Tabla 3.4: Medidas de tendencia central y dispersión de mujeres embarazadas de la muestra. . . .	24
Tabla 3.5: Correlación de Pearson entre variables continuas.	24

CAPÍTULO 1. INTRODUCCIÓN

1.1 MOTIVACIÓN

El cuerpo humano está conformado por distintos sistemas que hacen de sí un ente completo y complejo con un organismo en equilibrio. Uno de estos sistemas es el *Sistema Endocrino* que tiene la función principal de regular la producción de hormonas, las cuales se encargan de influir en la mayoría de las funciones del organismo, por ejemplo: funcionamiento de órganos, control de las distintas funciones del organismo, autorregulación, comportamientos del individuo, control de homeostasis, entre otras funciones. Así el sistema endocrino posee órganos que ayudan a formar el sistema como tal, uno de estos órganos es la llamada *Tiroides*, la cual es la encargada de producir hormonas tiroideas que ayudan al cuerpo utilizar energía, mantener la temperatura corporal y a que el cerebro, el corazón, los músculos y otros órganos funcionen normalmente [1]. No siempre nuestros sistemas y órganos funcionan como corresponde, en el caso de la Tiroides uno de los problemas más comunes son el *Hipotiroidismo* e *Hipertiroidismo* que tiene que ver principalmente con el desequilibrio de producción de *hormonas tiroideas*, en esta oportunidad el tema de interés se centra en el Hipotiroidismo, que se ocurre cuando la tiroides no es capaz de producir suficientes hormonas tiroideas T3 y T4. Esto último es un tema con gran potencial de estudio, ya que la población actual gradualmente ha incrementado este tipo de problemas, por lo cual se busca tener medicina suficientemente apta para dar soluciones eficientes a estos problemas, para esto se tiene evidencia de distintos datos de pacientes que nos pueden permitir determinar patrones de comportamientos y causas de tal enfermedad, lo que puede concluir en una gran aporte a la medicina para determinar causas de estos fallos en el organismo.

1.2 OBJETIVOS

- Describir el conjunto de datos de pacientes (Base de datos) a cabalidad de modo de tener conocimiento de todas las variables involucradas en el problema.
- Realizar un análisis estadístico a las variables de la base de datos.

1.3 ORGANIZACIÓN DEL DOCUMENTO

El documento consta de tres secciones: Descripción del problema, donde se entrega información detallada de la base de datos utilizada, fuentes, características de interés, entre otros. También se presenta una sección de análisis estadístico para las distintas variables que están presentes en la base de datos y por último una conclusión del trabajo realizado en este documento.

CAPÍTULO 2. DESCRIPCIÓN DEL PROBLEMA

Como bien se mencionó en el capítulo anterior el Hipotiroidismo es una falla que se produce en el sistema endocrino cuando la tiroides secreta menos hormonas de las que debería, pero lo anterior es una definición solo a grandes rasgos, por lo que esta es la oportunidad para profundizar más acerca del sistema endocrino, la tiroides y el problema del Hipotiroidismo con las hormonas de interés que se producen en bajos niveles:

- **¿Que es el sistema endocrino?:** Formalmente, el sistema endocrino está constituido por órganos y tejidos que segregan hormonas, sustancias químicas que funcionan como señales y se transportan, bien a través de la circulación sanguínea, o simplemente liberándose en el líquido intersticial y que afectan las células cercanas [2].
- **¿Que es la tiroides?:** La tiroides es una glándula que se ubica en la parte delantera del cuello, cercana a la manzana de adán. Su trabajo es formar hormonas tiroideas, volcarlas al torrente sanguíneo y entregarlas a todos los tejidos del cuerpo [1]. Esta glándula secreta la hormona conocida como *Tiroxina* (T4), compuesta por 4 átomos de yodo, también esta misma para cumplir otras funciones muta su estructura a 3 átomos de yodo y se conoce como *triiodotiroxina* (T3).

La tiroides, también es regulada por una hormona producida por la glándula pituitaria, esta hormona tiene el nombre de *Hormona estimulante de la tiroides* (TSH), los niveles de TSH regulan la producción de T4; a menores niveles de T4 la glándula pituitaria produce una mayor cantidad de TSH para estimular la tiroides a la producción de T4, a continuación se presenta un recurso gráfico de la relación entre T4 y TSH junto con la causante de condición de la tiroides.

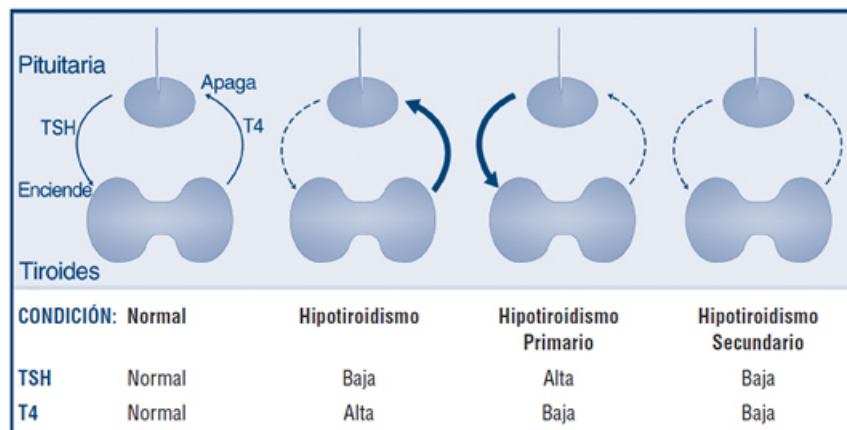


Figura 2.1: Condición de tiroides según relación T4-TSH [1].

2.1 DESCRIPCIÓN DE LA BASE DE DATOS

La base de datos para el análisis del problema corresponde a la base de datos proporcionada por el Garvan Institute of Medical Research, un instituto ubicado en Sidney, Australia, que gracias a Ross Quinlan en el año 1987, se recabaron datos de la enfermedad de tiroides de distintos pacientes.

Esta base de datos posee 2800 observaciones con 29 atributos que representan diversas características de interés para evaluar la condición de tiroides de un paciente, estos pueden ser continuos o booleanos. Además, esta base de datos posee una característica de clasificación para cada observación en cuanto al hipotiroidismo: Negativo, Hipotiroidismo Compensado, Hipotiroidismo Secundario, Hipotiroidismo Primario. La base de datos consta de dos archivos .data y .name, que corresponden respectivamente a la base de datos como tal y la caracterización de headers con descripciones de interés de la base de datos.

2.1.1. Descripción de clases y variables

1. **Age:** Indica la edad del paciente.
2. **Sex:** Indica si el paciente es hombre o mujer, toma valores M o F.
3. **On thyroxine:** Variable que toma valores booleanos, muestra si el paciente está o no en algún tratamiento con medicación de tiroxina. La tiroxina es la hormona elaborada por la glándula tiroides y que contiene yodo. La tiroxina aumenta la tasa de reacciones químicas en las células y ayuda a controlar el crecimiento y el desarrollo. La tiroxina también se puede producir en el laboratorio y se usa para tratar trastornos tiroideos. También se llama L-3,5,5'-tetrayodotironina, T4, y tetrayodotironina [3].
4. **Query on thyroxine:** Indica si el paciente se encuentra activo en una consulta sobre sus niveles de tiroxina, la variable toma valores booleanos.
5. **On antithyroid medication:** Indica si el paciente se encuentra en algún tratamiento con medicamentos que inhiben las funciones de hormonas tiroideas.
6. **Sick:** Indica si el paciente se encuentra enfermo. La variable toma valores booleanos.
7. **Pregnant:** Si el paciente es mujer, indica si esta se encuentra embarazada. La variable toma valores booleanos.
8. **Thyroid surgery:** Indica si el paciente ha pasado por cirugías que implique la tiroides. La variable toma valores booleanos.
9. **I131 treatment:** Indica si el paciente se encuentra bajo el tratamiento I131, este tratamiento emplea procedimientos de la medicina nuclear para personas diagnosticadas con hipertiroidismo, que consiste en ingerir dosis del isótopo radioactivo yodo I-131, el cual emite radiación. Este se absorbe en la

sangre para luego concentrarse en la glándula de la tiroides para comenzar a destruir las células de esta [4].

10. **Query hypothyroid:** Indica si el paciente está en consulta por hipotiroidismo. La variable toma valores booleanos.
11. **Query hyperthyroid:** Indica si el paciente está en consulta por hipertiroidismo. La variable toma valores booleanos.
12. **Lithium:** Indica si el paciente se encuentra en algún tratamiento con litio, la variable toma valores booleanos. El litio se utiliza para tratar y prevenir los episodios de manía (ánimo frenético, anormalmente emocionado) en las personas con trastorno bipolar (trastorno maníaco-depresivo; una enfermedad que provoca episodios de depresión, episodios de manía y otros estados de ánimo anormales). El litio se encuentra en una clase de medicamentos llamados agentes antimaníacos. Funciona al reducir la actividad anormal en el cerebro [5].
13. **Goitre:** Indica si el paciente padece de bocio. El término “bocio” simplemente se refiere a un agrandamiento anormal de la glándula tiroides. Es importante saber que la presencia de un bocio no necesariamente indica que la glándula tiroides está funcionando mal. El bocio puede ocurrir en una glándula que está produciendo demasiada hormona (hipertiroidismo), muy poca hormona (hipotiroidismo) o la cantidad correcta (eutiroidismo). El bocio indica que hay una condición que está causando un crecimiento anormal de la tiroides [6]. La variable toma valores booleanos.
14. **Tumor:** Indica si el paciente posee algún tipo de tumor (benigno o maligno) en la glándula tiroides. La variable puede tomar valores booleanos.
15. **Hypopituitary:** Indica si el paciente padece de hipopituitarismo. El hipopituitarismo es una afección en la cual la hipófisis no produce cantidades normales de algunas o de todas sus hormonas. En el hipopituitarismo, hay una falta de una o más hormonas hipofisarias. La falta de la hormona lleva a una pérdida de la función en la glándula u órgano que controla. Por ejemplo, la pérdida de la TSH de la tiroides ocasiona una pérdida de la función normal de la glándula tiroides [7]. La variable toma valores booleanos.
16. **Psych:** Indica si el paciente se encuentra en algún tratamiento psiquiátrico, que frecuentemente implica el uso de medicamentos. La variable toma valores booleanos.
17. **TSH measured:** Indica si el paciente ha sido sometido a pruebas para medir TSH. La prueba de TSH es la que determina el hipotiroidismo, hipertiroidismo o la normalidad de la glándula: Un nivel de TSH elevado indica que la glándula tiroides está fallando debido a un problema que afecta directamente a la glándula (hipotiroidismo primario). La situación opuesta, en la cual el nivel de TSH está bajo, generalmente indica que la persona tiene una glándula hiperactiva que está produciendo demasiada hormona tiroidea (hipertiroidismo) [1]. La variable toma valores booleanos.

18. **TSH:** Medida de los niveles de TSH u hormona estimulante de la tiroides. La variable toma valores continuos.
19. **T3 measured:** Indica si el paciente ha sido sometido a pruebas para medir la hormona triiodotiroxina (T3). La variable toma valores booleanos.
20. **T3:** Medida de los niveles de triiodotiroxina (T3). La variable toma valores continuos.
21. **TT4 measured:** Indica si el paciente ha sido sometido a mediciones de T4. TT4 significa Total T4, lo que implica que el examen consiste en determinar los niveles totales de tiroxina. La variable toma valores booleanos.
22. **T4:** Medida de los niveles de tiroxina (T4). La variable toma valores continuos.
23. **T4U measured:** Indica si el paciente ha sido sometido a pruebas de T4U. Más del 99 % de la T4 y T3 circulante en la sangre se une a una proteína llamada TBG (globulina fijadora de tiroxina, siglas en inglés). Estas hormonas vinculadas a la TBG son inocuas y no pueden ser utilizadas por los órganos y tejidos [8], a estas hormonas tiroideas se les denomina T4U. La variable toma valores booleanos.
24. **T4U:** Medida de los niveles de T4U. La variable toma valores continuos.
25. **FTI measured:** Indica si el paciente ha sido sometido a mediciones de FTI. La fracción de T4 libre (FTI) es la más importante para determinar cómo está funcionando la tiroides, y las pruebas que miden esta fracción se llaman T4 libre (FT4) y el Índice de T4 libre (FT4I o FTI). Las personas con hipertiroidismo tendrán FT4 o FTI elevados, mientras que los pacientes con hipotiroidismo tendrán un nivel bajo de FT4 o FTI [1]. La variable toma valores booleanos.
26. **FTI:** Medida de los niveles de FTI. La variable toma valores continuos.
27. **TBG measured:** Indica si el paciente ha sido sometido a mediciones de TBG. La globulina fijadora de tiroxina (TBG), es una glucoproteína que se une en la circulación sanguínea a las hormonas tiroideas T4 y T3, el análisis de TBG en la sangre pueden diagnosticar estado de normalidad de tiroides [9]. La variable toma valores booleanos.
28. **TBG:** Medida de los niveles de TBG. La variable toma valores continuos.
29. **Referral source:** Fuentes que se utilizaron para conformar la base de datos.

CAPÍTULO 3. ANÁLISIS ESTADÍSTICO E INFERENCIAL

Previo a realizar un análisis concreto de la base de datos, es imperioso hacer una barrida de limpieza sobre esta base de datos, este proceso se utiliza principalmente para solucionar problemas de datos incorrectos, incompletos, inexactos o irrelevantes, para luego modificarlos, sustituirlos o eliminarlos, dado que esto datos falsos o erróneos pueden llevar a conclusiones falsas.

El dataset utilizado de hipotiroidismo posee 2800 sujetos, existiendo varios sujetos que tienen datos incompletos, es decir, hubo algunos exámenes que no se les realizaron. Se decide en contra de la imputación, dado que esta asume que los datos son aleatorios, pero en este caso es más probable que el doctor decida realizar los exámenes sólo cuando el sujeto presente síntomas de hipotiroidismo. Si decide lo contrario podría significar que el sujeto se encuentra saludable (o presenta otro padecimiento), efecto que sería ocultado por la imputación de datos. Por estas razones se decide eliminar los pacientes que tienen datos incompletos.

Hay que notar dos casos especiales, el primero es la existencia de una variable que en su totalidad se encuentra incompleta, ninguno de los 2800 sujetos fue sometido al examen de TBG, por lo que se elimina esta variable (junto a su par TBG measured) de la base de datos. El segundo caso especial es el de un sujeto que dice tener 455 años en la tabla de datos, lo que es muy improbable por lo que se elimina a este sujeto del análisis. Luego de borrar los pacientes con datos incompletos o erróneos, también se eliminan las variables que indican si es que se realizaron tales exámenes o no, ya que todos los sujetos que quedan efectivamente se realizaron todos los exámenes. Finalmente, de los 2800 pacientes originales quedan 1946, es decir, el 69.5 % de la muestra total.

En la Figura 3.1 se muestra cómo se encuentran clasificados los 1946 pacientes de acuerdo a su diagnóstico.

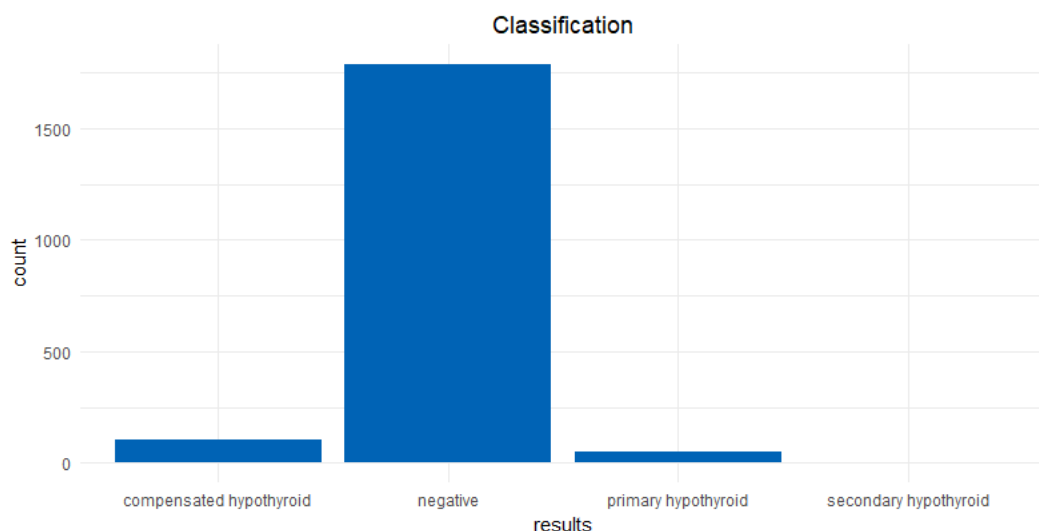


Figura 3.1: Cantidad de pacientes de acuerdo a su diagnóstico.

3.1. Análisis de variables booleanas

3.1.1. Género

Primero se debe tener en mente que la cantidad promedio de hormonas tiroideas varía dependiendo del sexo del paciente. A simple vista en la Figura 3.2 se podría decir que las mujeres presentan más casos de hipotiroidismo que los hombres, pero esto no considera que en la muestra la cantidad de mujeres es 1.92 veces mayor que la cantidad de hombres. Viéndolo en porcentajes 5.8 % de los hombres padecen de hipotiroidismo y 9.2 % de las mujeres padecen de hipotiroidismo, por lo que, al menos en la muestra dada, para las mujeres si es más probable padecer de hipotiroidismo.

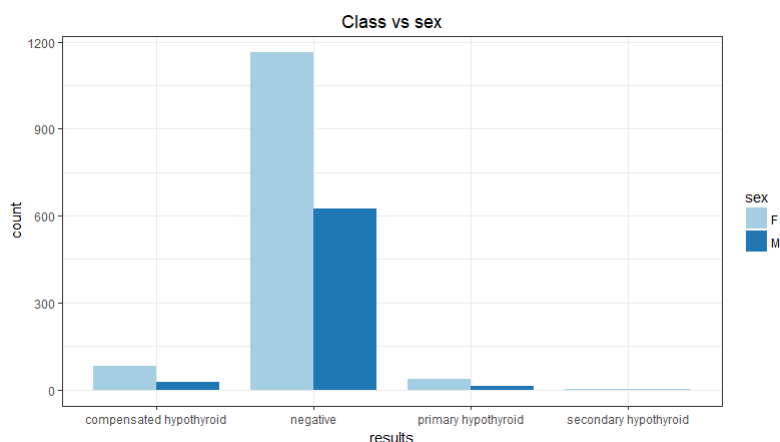


Figura 3.2: Resultados del diagnóstico de acuerdo al sexo de los pacientes.

3.1.2. Embarazo

Un embarazo normal trae consigo una serie de cambios fisiológicos y hormonales que alteran la función tiroidea. Esto significa que las pruebas de laboratorio de la función tiroidea deben ser interpretadas cuidadosamente durante el embarazo [10]. Se debe tener en cuenta este cambio en las hormonas tiroideas, pero mirando la Figura 3.3 se puede apreciar que el embarazo no es un factor clave en el análisis.

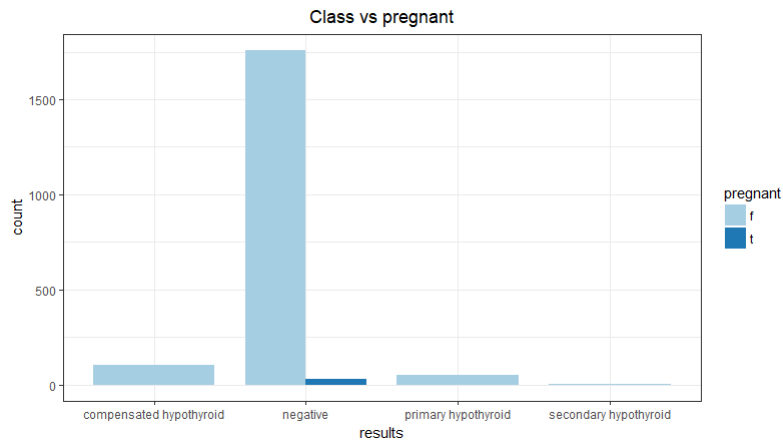


Figura 3.3: Resultados del diagnóstico de acuerdo a si la paciente se encuentra embarazada.

3.1.3. Litio

En dosis pequeñas este medicamento puede bloquear la secreción de hormonas tiroideas, y su uso prolongado o excesivo podría inhibir la captación de Iodo por la glándula tiroides [11], es decir, podría inducir hipotiroidismo. En la Figura 3.4 se aprecia que los pacientes que han utilizado este medicamento son muy pocos en comparación a los no lo han utilizado.

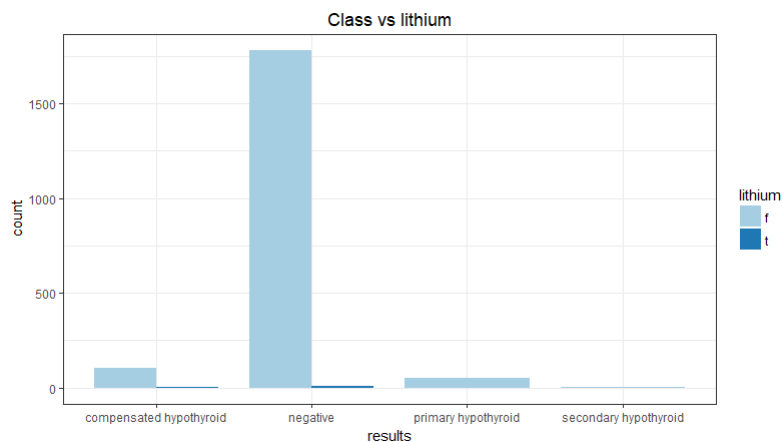


Figura 3.4: Resultados del diagnóstico de acuerdo a si el paciente ha sido medicado con litio.

Como no se puede ver muy bien en la Figura 3.4 si es que hay alguna diferencia significativa entre los pacientes que han utilizado litio y que luego han sido diagnosticados algún tipo de hipotiroidismo, se utiliza la Figura 3.5, donde sólo se muestran los pacientes que han sido tratados con Litio (para las próximas variables es utilizará este mismo tipo de gráfico). Hay 10 pacientes que tuvieron resultado negativo (salud) y 1 paciente con resultado positivo (hipotiroidismo), si es que hubiera más pacientes que hayan sido tratados con litio y se mantuviera la misma proporción se podría llegar a la conclusión de que el 9 % de los pacientes que fueron tratados con algún medicamento de litio padecerán de hipotiroidismo.

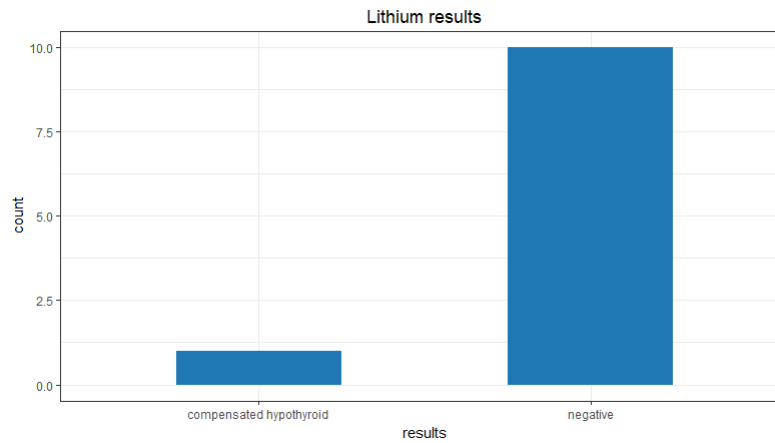


Figura 3.5: Resultados del diagnóstico para sólo los pacientes que han sido medicados con litio.

3.1.4. Operado de la tiroides

Para el análisis del hipotiroidismo se puede ver en la Figura 3.6 que esta variable no presenta información significativa como para convertirse en un factor importante en la detección del hipotiroidismo. Hay 24 pacientes saludables y sólo 1 que padece de hipotiroidismo (4 % de los pacientes que fueron sometidos a una cirugía de la tiroides, ahora padecen de hipotiroidismo, los mismos problemas que en la sección anterior se aplican).

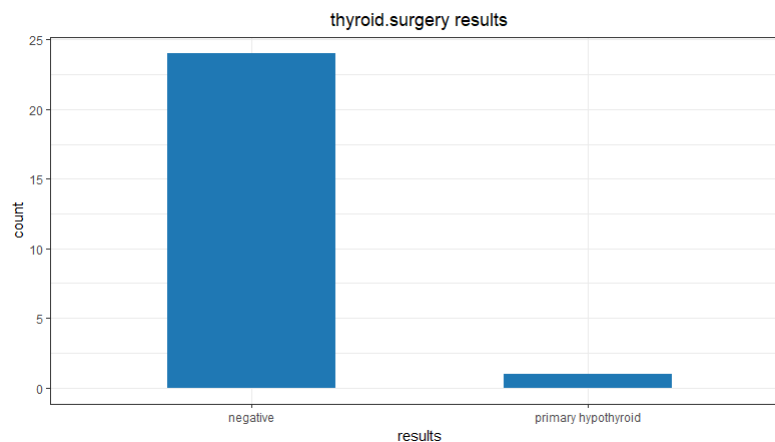


Figura 3.6: Resultados del diagnóstico dado que el paciente ha pasado por cirugías que implique la tiroides.

3.1.5. Psych

En la Figura 3.7 se puede apreciar que hay 123 pacientes saludables y 6 pacientes que ahora han sido diagnosticados con hipotiroidismo compensado, lo que sigue siendo una cantidad baja, pero mucho mayor que la cantidad de pacientes en las variables Litio y los que fueron operados en la tiroides. La conclusión que se podría sacar es que 4.6 % de los pacientes que se encuentran en algún tratamiento psiquiátrico, ahora padecen de hipotiroidismo compensado.

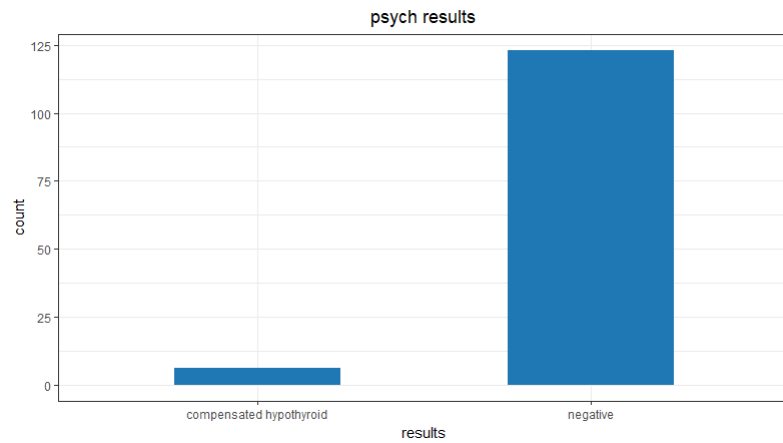


Figura 3.7: Resultados del diagnóstico dado que el paciente se encuentre en algún tratamiento psiquiátrico.

3.1.6. Bocio (Goitre)

El Bocio se refiere a un agrandamiento anormal de la glándula tiroides, esto podría llevar a un pensamiento incorrecto de que tal agrandamiento sea dañino, pero como se puede apreciar en la Figura 3.8, este agrandamiento al menos no le causó hipotiroidismo a ninguno de los 1946 pacientes de la muestra.

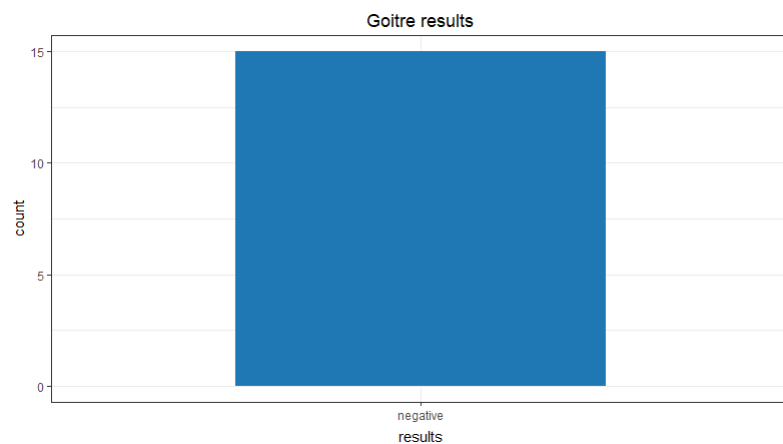


Figura 3.8: Resultados del diagnóstico dado que el paciente padezca de bocio.

3.1.7. Tumor

La Figura 3.9 muestra el efecto que tiene un tumor en ser la posible causa de hipotiroidismo, hay 40 pacientes saludables y 4 pacientes que sufren de hipotiroidismo. La cantidad de pacientes le da más credibilidad a esta Figura en comparación con los pacientes que son tratados con Litio donde hay sólo 1 paciente con hipotiroidismo, Figura 3.5.

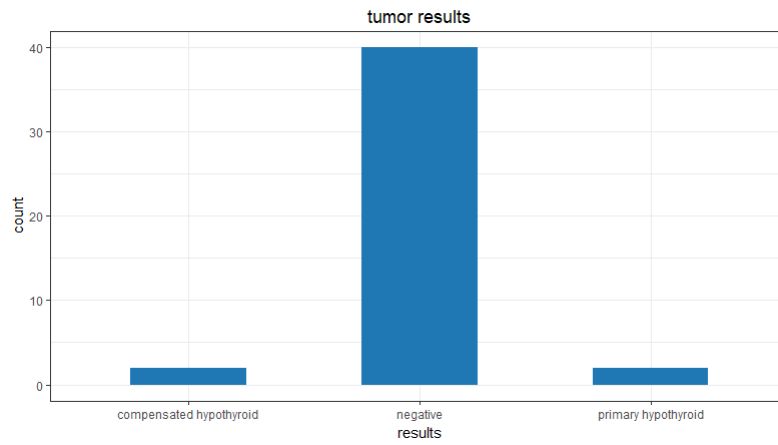


Figura 3.9: Resultados del diagnóstico dado que el paciente posea algún tipo de tumor en la glándula tiroides.

3.2. Análisis de variables continuas

3.2.1. Medidas de tendencia central y dispersión

Para comprender mejor el comportamiento y el efecto que tienen estas variables en el hipotiroidismo, se comienza por confeccionar la siguiente tabla:

Tabla 3.1: Medidas de tendencia central y dispersión de la muestra.

Variable	Media	Mediana	Moda	Desviación
Edad	53.00976	55	59	18.79502
TSH	4.37778	1.3	0.2	18.43231
T3	2.01490	2	2	0.82493
TT4	108.7557	104	101	35.45565
T4U	0.99862	0.98	0.99	0.19905
FTI	110.0048	107	96	32.3947

Edad: Aunque parezca que la moda debiera ser 70 aproximadamente, hay 59 pacientes de 59 años y 58 de 70 años. La Figura 3.10 muestra que el 50 % de las personas que pasan por estos exámenes son mayores a 55 años. Gracias a este gráfico se logra detectar que existía un sujeto en la base de datos que tenía 455 años, sujeto que ya fue eliminado de la muestra.

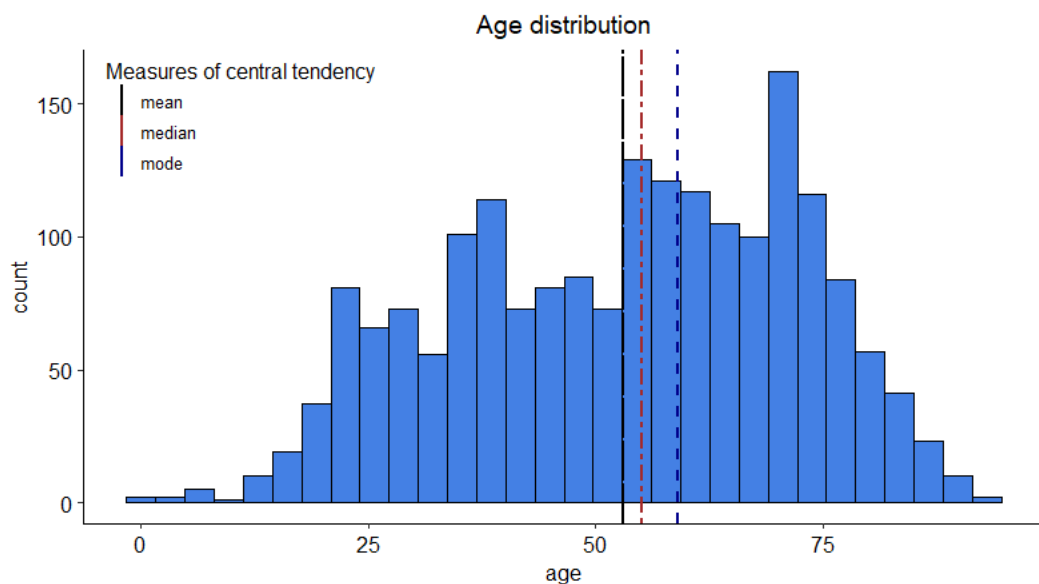


Figura 3.10: Distribución y medidas de tendencia central de la edad.

TSH: Según [12], el rango normal de TSH puede variar dependiendo de la edad del paciente, el género del paciente y el embarazo. Será considerado el rango sugerido de 0.45 hasta 4.12 mUI/L como el rango normal de TSH, valores mayores pueden ser indicador de hipotiroidismo. En la Figura 3.11 se puede ver que hay varios sujetos que tienen valores de TSH extremadamente altos, ~ 100 veces sobre el rango normal.

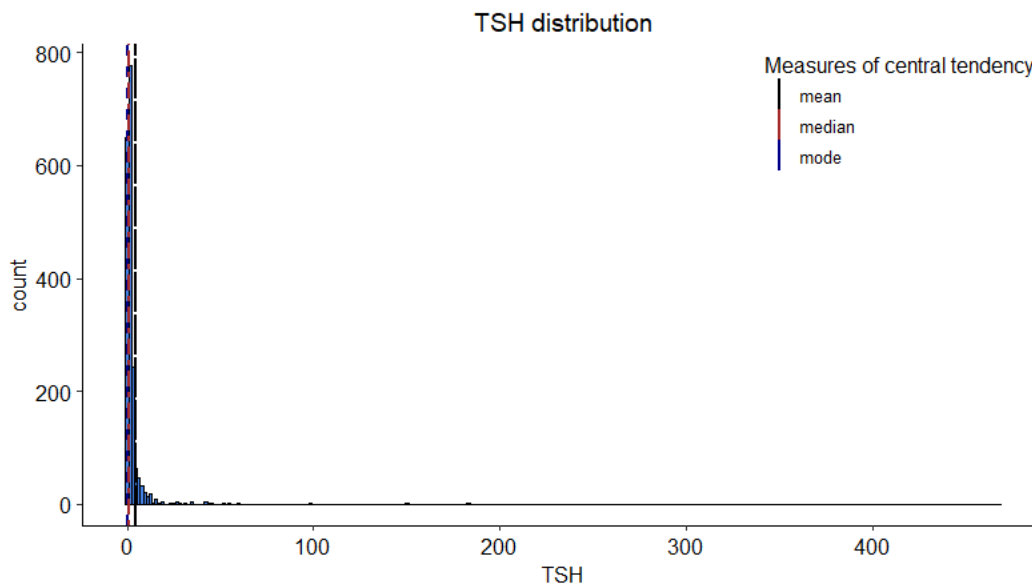


Figura 3.11: Distribución y medidas de tendencia central de TSH.

Para poder analizar mejor la distribución de TSH, se utiliza la Figura 3.12, donde se ignoran los datos muy altos. El promedio es de 4.37778, valor por sobre el rango normal. La mediana es igual a 1.3, mientras que el tercer cuartil es igual a 2.5, lo que significa que el 75 % de los pacientes tiene poco riesgo.

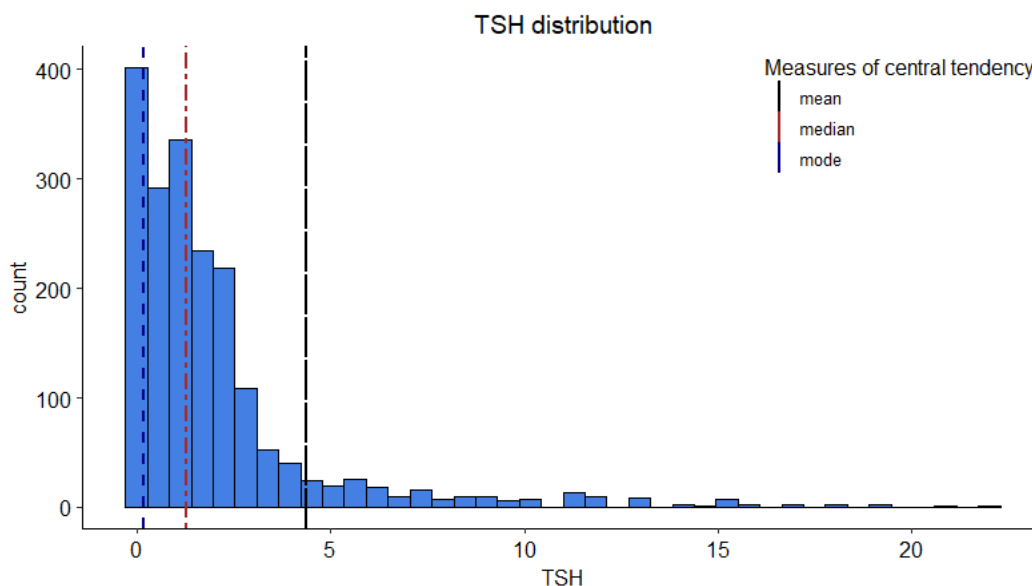


Figura 3.12: Acercamiento sin cola a la derecha de la distribución y medidas de tendencia central de TSH.

T3: El rango normal de T3 es de 1.2 hasta 2.6 nmol/L [13]. Según Figura 3.13 y Tabla 3.1 en promedio los pacientes tienen 2.0 nmol/L de T3, por lo que se encuentran en un rango normal, lo mismo con la mediana y la moda (2 nmol/L de T3).

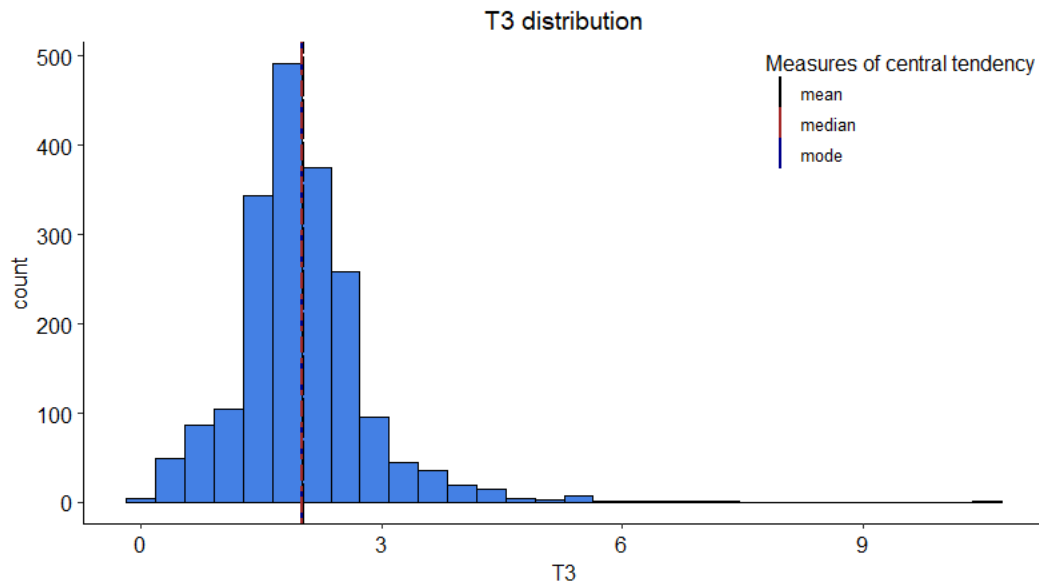


Figura 3.13: Distribución y medidas de tendencia central de T3.

TT4: El rango normal del total de tiroxina es de 64.4 hasta 161 nmol/L [13]. Según Figura 3.14 y Tabla 3.1 en promedio los pacientes tienen 108.7557 nmol/L de tiroxina, por lo que se encuentran en un rango normal, lo mismo con la mediana y la moda (104 y 101 nmol/L de T4 respectivamente).

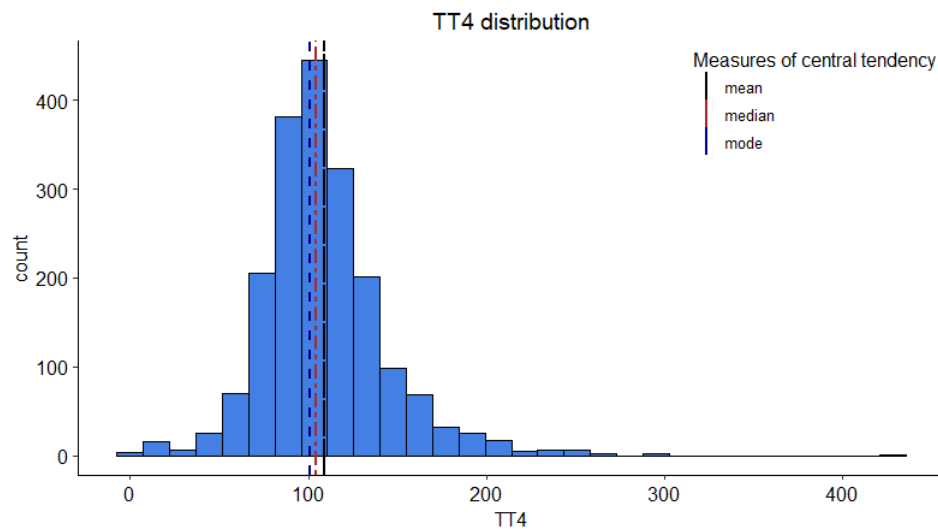


Figura 3.14: Distribución y medidas de tendencia central de TT4.

T4U: Para esta hormona tiroidea, valores cercanos al 1 ng/dl indican salud en el paciente. En promedio los pacientes tienen 0.99862 ng/dl de T4U, lo cual indica que en promedio los pacientes se encuentran saludables en cuanto a T4U. La moda y mediana también se encuentran dentro del rango, siendo sus valores 0.99 y 0.98 respectivamente.

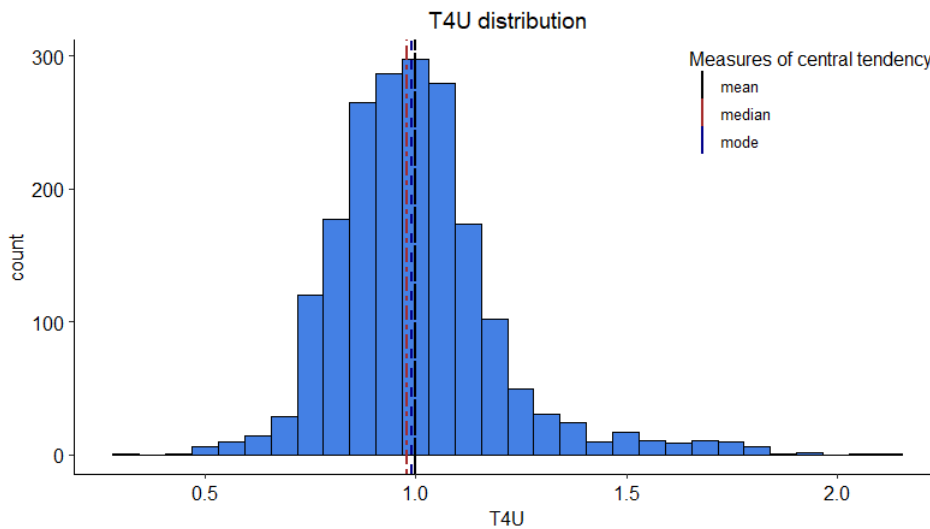


Figura 3.15: Distribución y medidas de tendencia central de T4U.

FTI: El rango normal del índice de tiroxina libre en un sujeto es de 34.6 hasta 121 nmol/L [13]. Según Figura 3.14 y Tabla 3.1 en promedio los pacientes tienen 110.0048 nmol/L en este índice, por lo que se encuentran en un rango normal, lo mismo con la mediana y la moda (107 y 96 nmol/L de T4 respectivamente). Es importante notar que mientras más bajo este valor, mayor riesgo tiene un paciente de padecer de hipotiroidismo.

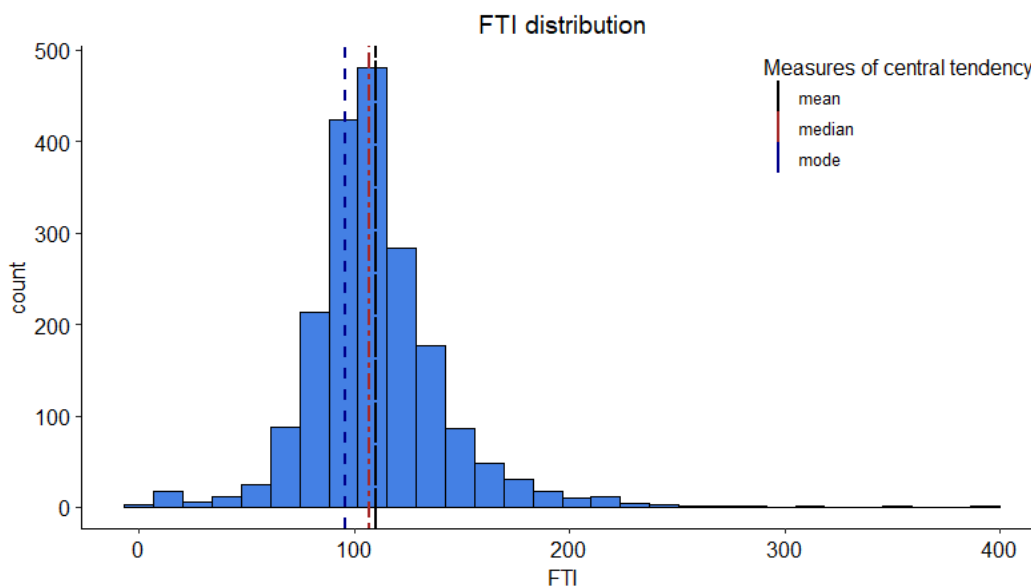


Figura 3.16: Distribución y medidas de tendencia central de FTI.

Dado que el rango normal de hormonas tiroideas varía de acuerdo al género y embarazo [13], se confeccionan las siguientes tablas para mostrar las medidas de tendencia central y dispersión al separar la muestra de acuerdo al género de los pacientes, separando a las mujeres en embarazadas y no embarazadas.

Tabla 3.2: Medidas de tendencia central y dispersión de hombres de la muestra.

Variable	Media	Mediana	Moda	Desviación
Edad	52.56992	55	59	18.09509
TSH	3.35272	1.3	1.3	12.91266
T3	1.93263	2	2	0.76207
TT4	100.8042	99	86	27.47684
T4U	0.93072	0.92	0.91	0.14920
FTI	109.1368	107	105	27.77504

Tabla 3.3: Medidas de tendencia central y dispersión de mujeres no embarazadas de la muestra.

Variable	Media	Mediana	Moda	Desviación
Edad	53.76141	57	70	19.05422
TSH	5.01614	1.4	0.2	20.96237
T3	2.02466	2	1.8	0.83168
TT4	111.6508	107	103	37.34526
T4U	1.01975	1	0.99	0.19177
FTI	110.704	107	100	34.80871

Tabla 3.4: Medidas de tendencia central y dispersión de mujeres embarazadas de la muestra.

Variable	Media	Mediana	Moda	Desviación
Edad	32.8125	32.5	34	9.28209
TSH	0.76359	0.2	0.02	1.58024
T3	3.34375	3.2	2.5	0.66329
TT4	161	152	139	45.20848
T4U	1.585	1.64	1.55	0.23820
FTI	100.75	97.5	104	20.38659

3.2.2. Relación entre variables

Para entender mejor el efecto de estas variables, además de la distribución individual de cada variable, se pueden explorar las relaciones que tienen estas variables continuas entre sí, utilizando la correlación de Pearson para confeccionar la siguiente matriz de correlación.

Tabla 3.5: Correlación de Pearson entre variables continuas.

Variable	Edad	TSH	T3	TT4	T4U	FTI
Edad	1.00000	-	-	-	-	-
TSH	-0.03157	1.00000	-	-	-	-
T3	-0.28399	-0.18042	1.00000	-	-	-
TT4	-0.08570	-0.28825	0.58184	1.00000	-	-
T4U	-0.19589	0.05597	0.46573	0.45049	1.00000	-
FTI	0.03578	-0.33108	0.34687	0.78637	-0.17076	1.00000

Relaciones importantes que se pueden ver en la matriz de correlación (Tabla 3.5):

- Edad tiene una relación inversa con todas las variables excepto FTI, esto se debe a que mientras el sujeto envejece, produce menos hormonas tiroideas y tiene mayor riesgo de padecer alguna enfermedad relacionada a la tiroides.
- Mientras haya más hormonas estimulantes de la tiroides (TSH) menor será la cantidad total de tiroxina (TT4) y T3 en el paciente. Esto se debe a que, en los pacientes con hipotiroidismo, la hipófisis necesita mantener niveles de TSH más elevados que lo habitual con el fin de estimular constantemente la tiroides a aumentar su producción de T3 y T4.
- Hay una relación directa entre TT4 y FTI, esto se debe a que el índice de tiroxina libre se puede calcular utilizando el TT4 [13], por lo que sería extraño que no hubiera una relación entre estas variables.

CAPÍTULO 4. CONCLUSIONES

La base de datos utilizadas es una importante evidencia para verificar una de las enfermedades más comunes de los últimos tiempos. El capítulo de descripción del problema en el que se abordan características de la tiroides y su funcionamiento, además de la descripción de clases y variables de la base de datos, permitieron que se pudiera absorber conocimiento de este fenómeno biológico en detalle, de modo de contextualizar el problema no tan solo en el ámbito estadístico y de minería de datos, sino que también desde un punto de vista de la ciencia biológica.

Observando la muestra de datos a simple vista, hay varios datos incompletos, normalmente este tipo de anomalías se deben a errores al momento de recopilar la información, o alguna falla humana al momento de digitar los datos. Pero como este es un estudio médico, es más probable que lo sucedido sea que los pacientes que no se realizaron todos los exámenes fueron a causa de que el doctor a cargo se haya percatado de que el paciente no presentaba síntomas acordes a una disfunción de tiroides. Entonces, al realizar la limpieza de los datos, se decide eliminar todos estos sujetos con datos incompletos incluyendo un caso especial donde si hubo error humano y existía un paciente con 455 años. Finalmente se pierde el 30 % de la información debido a los datos incompletos, datos que de igual manera podrían ser útiles para ayudar en la predicción de pacientes que no padecen de hipotiroidismo si es que se hubieran realizado todos los exámenes.

Luego, por medio del análisis de las variables booleanas, se descubre que, al menos con la muestra dada, el Bocio no es causa de hipotiroidismo. Lo mismo ocurre con el embarazo de una mujer, de las 32 mujeres embarazadas que hay en la muestra, ninguna presenta algún tipo de hipotiroidismo. Además, el sexo del paciente pareciera influir levemente en si tal paciente padece de hipotiroidismo, dado que se obtuvo que el 5.8 % de los hombres de la muestra padecen de hipotiroidismo y el 9.2 % de las mujeres de la muestra padecen de hipotiroidismo.

Por otro lado, las variables como Litio, Psych, Tumor y si el paciente se ha operado de la tiroides parecieran tener una pequeña influencia en si el paciente padece de hipotiroidismo, pero no se puede inferir nada concreto por medio del análisis realizado.

El análisis de las variables continuas sirve para encontrar pacientes que podrían estar en riesgo de sufrir de hipotiroidismo, dado que tienen niveles fuera del rango normal, específicamente en los niveles de TSH que poseen algunos de los pacientes de la muestra, estos llegan a ser más que ~ 10 veces más altos que el máximo del rango normal de TSH sugerido por la literatura [12].

Sumando el análisis de las relaciones entre las variables continuas, se puede verificar la teoría, dado que mientras más edad tiene el paciente, menos hormonas tiroideas (TSH, T3 y TT4) produce. Mientras mayor el nivel de TSH menos tiroxina tendrá el paciente. También con el índice de tiroxina libre (FTI) el cuál se calcula utilizando TT4 y como se debiera esperar, posee una relación directa con esta hormona.

CAPÍTULO 5. BIBLIOGRAFÍA

- [1] A. T. Association, *Pruebas De Función Tiroidea*, 2016. dirección: <https://www.thyroid.org/las-pruebas-de-funcion-tiroidea/>.
- [2] M. I. C. Gonzalez, “Sistemas y aparatos del organismo(III). Topografía corporal”, en *Fisiopatología General*, ép. ssues, vol. 1, Madrid, España: Ediciones Paraninfo, 2016, págs. 117-121, ISBN: 9788428337984.
- [3] I. nacional del cáncer, *Definición tiroxina*, 2017. dirección: <https://www.cancer.gov/espanol/publicaciones/diccionario/def/tiroxina>.
- [4] R. S. of North America, *Radioactive Iodine (I-131) Therapy*, 2018. dirección: <https://www.radiologyinfo.org/en/info.cfm?pg=radioiodine>.
- [5] —, *Litio*, 2018. dirección: <https://medlineplus.gov/spanish/druginfo/meds/a681039-es.html>.
- [6] A. T. Association, *Bocio*, 2016. dirección: <https://www.thyroid.org/bocio/>.
- [7] B. nacional de medicina de los Estados Unidos, *Hipopituitarismo*, 2018. dirección: <https://medlineplus.gov/spanish/ency/article/000343.htm>.
- [8] D. P. Pinheiro, *HORMONAS TSH Y T4 LIBRE – EXAMEN DE LA FUNCIÓN TIROIDEA*, 2018. dirección: <https://www.mdsaude.com/es/2015/12/tsh-t4-libre-tiroides.html>.
- [9] P. Li, O. E. Janssen, K. Takeda, R. H. Bertenshaw y S. Refetoff, “Complete thyroxine-binding globulin (TBG) deficiency caused by a single nucleotide deletion in the TBG gene”, *Metabolism-Clinical and Experimental*, vol. 40, n.º 11, págs. 1231-1234, 1991.
- [10] A. T. Association, *La Enfermedad De La Tiroides Y El Embarazo*, 2016. dirección: <https://www.thyroid.org/enfermedad-tiroides-embarazo/>.
- [11] D. S. Scarone, *¿Sabías que el litio afecta a la glándula tiroides?*, 2015. dirección: <http://tuendocrinologo.com/site/inicio/articulos-por-tema/endocrinologia/el-litio-afecta-a-la-tiroides.html>.
- [12] E. K. Luo, *All About Standard TSH Ranges by Age and Life Stage*, 2018. dirección: <https://www.healthline.com/health/tsh-normal-range-by-age>.
- [13] T. G. L. of Women’s Medicine, *Thyroid Function Tests*, 2014. dirección: https://www.glowm.com/lab_text/item/87.