



Universidad Autónoma De Nuevo León
Facultad de Ciencias Físico Matemáticas

Inferencia Estadística

Jesús Andrés Solís López

PIA

Grupo: 003

Integrantes:

Cinthyá Yesenia López Díaz 1658003

Marco Antonio Oviedo Acevedo 1851698

Zaira Tamara Escareño Loera 1868734

Arely Yazmín Reséndiz Pinales 1869432

Eliud Narváez moreno 1863955

Luis Adrián Navarro García 1866581

Yaziel Gibran Barbosa Alcocer 1851004

7 de noviembre de 2021

Índice

Índice	2
Introducción	3
Antecedentes teóricos y metodológicos del estudio.....	4
Planteamiento del Problema	6
Descripción de Variables.....	9
Estadísticos Básicos.....	12
Distribuciones y Medias Estadísticas.....	14
Visualización de los Datos	19
Prueba de Hipótesis	22
Ejercicios	25
Conclusión.....	26
Anexos	27
Bibliografía.....	32

Introducción

En la vida hay determinantes y preguntas que son necesarios saber para la toma de decisiones ya sea desde el ámbito académico hasta lo profesional, lo cual permite el desarrollo correcto de una empresa o el análisis de una población para tomar decisiones al respecto.

A través de la historia grandes matemáticos y físicos, crearon una ciencia que ayuda a determinar la muestra de población, márgenes de error, intervalos de confianza, entre muchos otros más; lo cual ha permitido que se determine la población de un país, aplicarse en la medicina a la hora de estudiar las enfermedades y hasta optimizar una empresa, mediante fórmulas matemáticas y modelos para realizar de manera ágil y rápida las diferentes operaciones planteadas, es por eso que este trabajo es de gran relevancia ya que se podrá obtener todo lo aprendido en este curso.

En este proyecto exploraremos los aspectos básicos del análisis estadístico y el por qué es importante utilizarlo en materias como la ciencia de datos. Utilizaremos herramientas de programación como el lenguaje de R y Python para analizar los datos de venta de la comunidad residencial de lujo dentro de Walt Disney World Resort en Lake Buena Vista, Florida, llamado Golden Oak. Tomaremos en cuenta distintas variables como la ciudad, renta, el estado, garantía hipotecaria entre muchas otras variables más que se definirá en el desarrollo del proyecto.

Antecedentes teóricos y metodológicos del estudio.

ESTADISTICA DESCRIPTIVA. El propósito de la estadística aplicada es el de obtener conclusiones de una población en estudio, examinando solamente una parte de ella denominada muestra.

ESTADISTICA INFERENCIAL. La estadística es una rama de las matemáticas encargada de reunir, organizar y analizar datos generalmente numéricos, ayuda a resolver problemas y además permite luego de realizados los cálculos tomar decisiones que puedan beneficiar al contexto que las estudia. La estadística y los procedimientos que con ella pueden realizarse han permitido de manera efectiva describir con exactitud datos de casi todas las ramas del conocimiento entre ellas: economía, psicología, política, física, biología, química, medicina e informática y ha servido como herramientas útiles para encontrarle relación a muchos datos estudiados por estas ciencias.

En la actualidad para un estadístico el trabajo va más allá de reunir datos y calcularlos, debe encargarse además de la difícil tarea de interpretar toda la información obtenida en los procesos estadísticos para que tenga un valor realmente importante.

La estadística se encuentra dividida en dos grandes ramas, cada una con un propósito específico:

- La estadística inferencial
- La estadística descriptiva

Nuestro estudio está basado en la estadística inferencial por lo que ampliaremos el concepto de esta en las siguientes secciones.

Marco Teórico.

La estadística inferencial es una parte de la estadística que solo trabaja con algunos de los datos de una población existente dentro de un grupo de elementos observados; es decir, solo toma una muestra n de los N elementos existente. Una vez que se obtiene este reducido grupo de datos la estadística inferencial trata de encontrar aspectos o propiedades relevantes para toda la población y basados en ellos tomar decisiones. Para obtener dichos resultados es necesario fundamentarse en cómo se selecciona la muestra, como realizar la inferencia de los datos y además la confianza que se puede tener en la información obtenida.

Cabe recalcar que para obtener datos fiables el nivel de conocimiento y comprensión de estadística, matemática y probabilidades debe de ser alto pues se debe recordar que los procedimientos están basados en pequeñas muestras las

cuales pueden sufrir variación. Con toda la información proporcionada es notorio que la estadística inferencial puede proveer de modelos importantes para estudiar un sinnúmero de datos multivalentes.

Métodos tales como componentes parciales, escalado, multidimensional, análisis de correspondencia, análisis de conglomerados, análisis factorial, análisis discriminante, entre otros, brindan a los estudiosos grandes posibilidades de entender y predecir el comportamiento que los datos pueden tomar dada una condición.

Empresa Golden Oak

Golden Oak en Walt Disney World Resort es una comunidad residencial de lujo dentro de Walt Disney World Resort en Lake Buena Vista, Florida. Fue diseñado por Walt Disney Imagineering, es propiedad y está operado por una subsidiaria de Disney recién formada, Golden Oak Realty. La primera fase de desarrollo se encuentra al sureste del Parque Temático Magic Kingdom en Bay Lake. El área fue nombrada para rendir homenaje al Golden Oak Ranch de Walt Disney en California.

Golden Oaks es un conjunto residencial inspirado en los famosos personajes de los cuentos. Sí, esos que han formado parte del imaginario colectivo por generaciones y que se niegan a morir.

Se trata de una experiencia de lujo llevada a sus límites. El precio de cada una de sus viviendas empieza a partir de los 1.8 millones de dólares; un monto elevado para la mayoría, pero adecuado para los servicios y amenidades que ofrece.

Cada casa puede ser confeccionada al gusto, pero ciertos parámetros deberán ser respetados; el más importante de ellos es que dentro de cada hogar se harán presentes los más mínimos detalles que nos recuerden que la ilusión se vive en Disney todos los días.

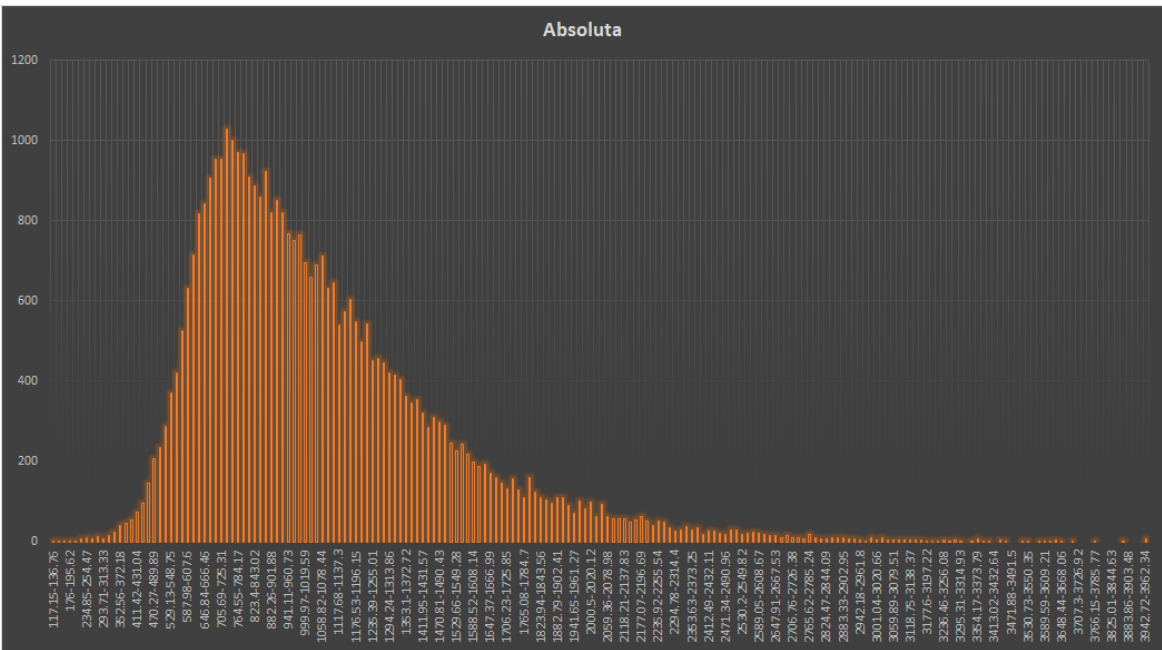
PLANTEAMIENTO DEL PROBLEMA

Como empresa queremos analizar la información de nuestros clientes para mejorar el servicio que ofrecemos, así como nuestras campañas de marketing para incrementar nuestras ventas, por lo que revisaremos los datos de nuestros compradores de agosto del 2017 de nuestras residencias de lujo Golden Oak. Decidimos empezar analizando los estados de vivienda donde nuestros clientes tienen una mayor compra, dado como resultado que los estados donde las personas compran más son de Florida con 2289 registros, New York con 2565 registros, Texas con 2767 registros y California con 4183 registros.



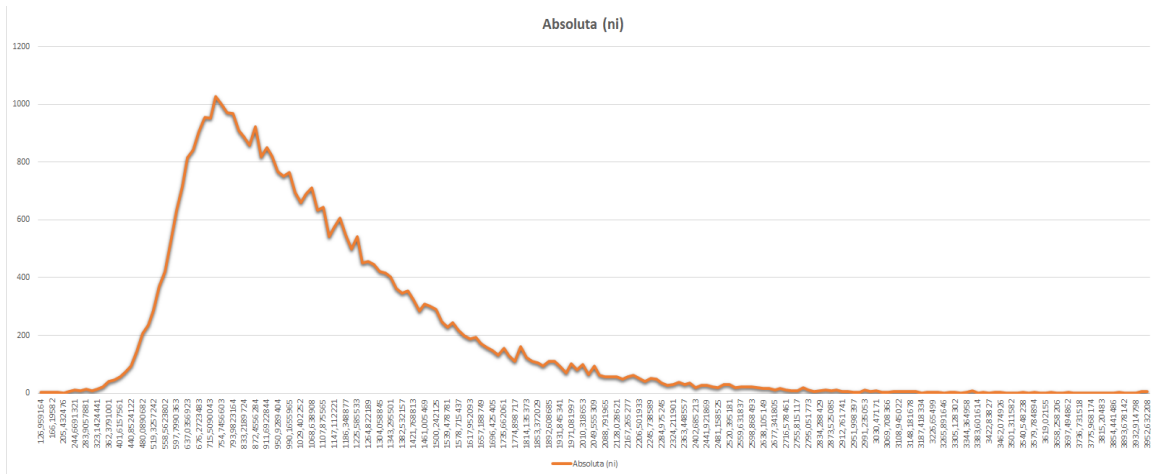
*Tabla de frecuencia de los estados en el apartado de anexos.

Renta Promedio (rent_mean): Histograma.

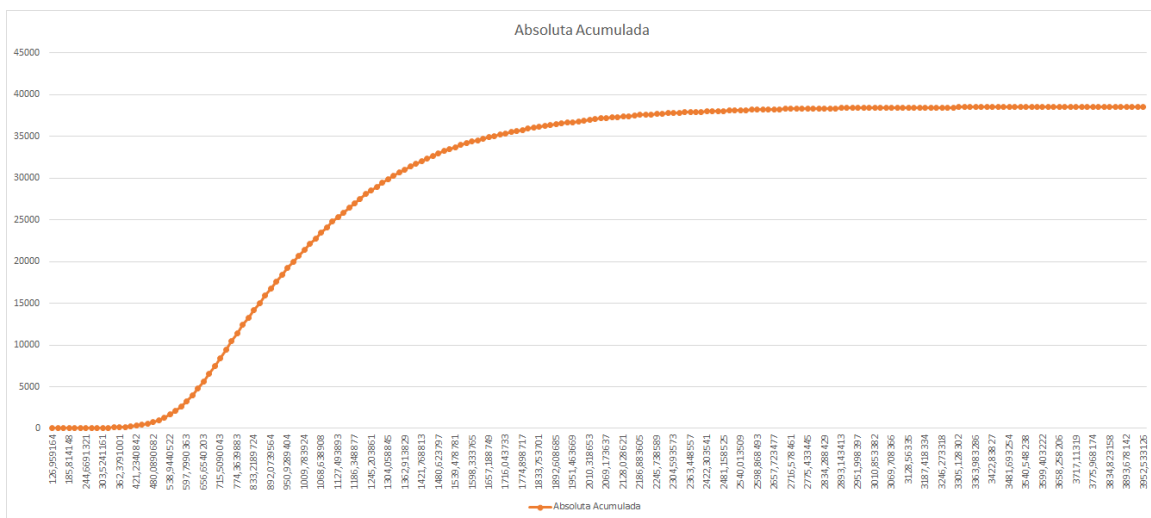


Esta gráfica muestra los intervalos y la frecuencia absoluta de las rentas promedio de los pagos de las residencias, podemos ver que nuestros clientes se decantan por las rentas de entre 725.31 a 744.93 con un promedio de 1028 registros.

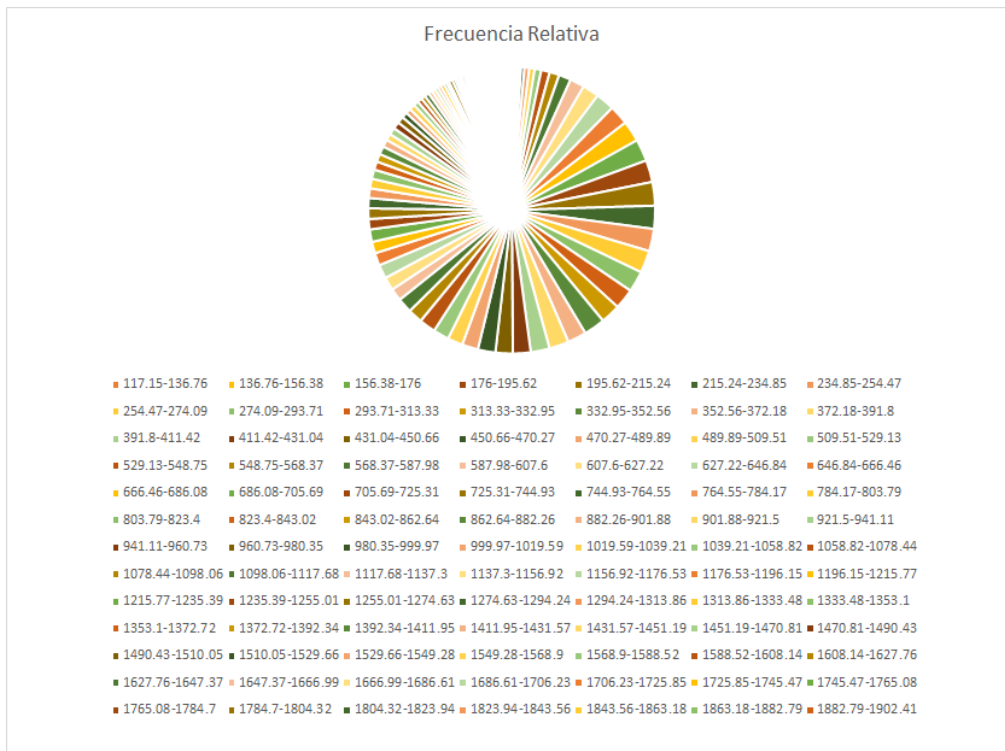
- Polígono de Frecuencia.



- Ojiva.



- Gráfico de Pastel.



*Tabla de frecuencia de la renta media (rent_mean) en el apartado de anexos.

DESCRIPCIÓN GENERAL DE VARIABLES

➤ **Costos mensuales de la hipoteca y del propietario**

Suma de pagos de hipotecas, escrituras de fideicomiso, contratos de compraventa o deudas similares sobre la propiedad, impuestos de bienes inmuebles, seguro contra incendios, riesgos e inundaciones en la propiedad, servicios públicos y combustibles, también incluye, en su caso, la tarifa mensual de condominio y costos de casas móviles (pagos a plazos de préstamos, impuestos a la propiedad personal, alquiler del sitio, derechos de registro y derechos de licencia). Los costos de propietario mensuales seleccionados se tabularon para todas las unidades ocupadas por el propietario, y por lo general se muestran por separado para las unidades "con una hipoteca" y para unidades "sin hipoteca".

➤ **Alquiler bruto**

Es el alquiler del contrato más el costo mensual promedio estimado de los servicios públicos (electricidad, gas, agua y alcantarillado) y combustibles (petróleo, carbón, queroseno, madera, etc.) que paga el arrendatario. El alquiler bruto está destinado a eliminar diferenciales que resultan de la inclusión de servicios públicos y combustibles como parte del pago del alquiler. Los costos estimados de agua, alcantarillado y combustibles se informan sobre una base de 12 meses, pero se convierten en cifras mensuales para las tabulaciones.

➤ **Ingresos del hogar y familiares**

Los ingresos del hogar incluyen la suma del jefe de hogar y todas las demás personas mayores de 15 años que residen en el hogar, sean parientes del jefe de hogar o no. Aunque las estadísticas de ingresos familiares cubren los últimos 12 meses, las características de los individuos y la composición de los hogares se refiere a la entrevista realizada con anterioridad. Así, los ingresos del hogar no incluyen las cantidades recibidas por personas que fueron miembros del hogar durante todo o parte de los últimos 12 meses o si estas personas no residían en el hogar en el momento de la entrevista.

DEFINICIONES DE CAMPO DE UBICACIÓN

- **UID:** ID de la ubicación analizada, ubicación de identificación compatible en todo todas las ubicaciones de investigación de Golden Oak.
- **STATITED:** Código de estado informado por la Oficina del Censo de EE. UU.

- **State:** Nombre del estado informado por la Oficina del Censo de EE. UU.
- **Pop:** Población masculina y femenina en el área.
- **Male_pop:** Población masculina en el área
- **Female_pop:** Población femenina en el área

DEFINICIONES DE CAMPO ESTADÍSTICO

- **second_mortgage:** Porcentaje de casas con una segunda hipoteca
- **home_equity:** Porcentaje de viviendas con un préstamo con garantía hipotecaria.
- **Deuda:** Porcentaje de viviendas con algún tipo de deuda.
- **second_mortgage_cdf:** Valor de distribución acumulado de uno menos el porcentaje de viviendas con segunda hipoteca. El valor se utiliza como característica de rendimiento.
- **home_equity_cdf:** Valor de distribución acumulado de uno menos el porcentaje de viviendas con
- préstamo con garantía hipotecaria. El valor se utiliza como característica de rendimiento.
- **deuda_cdf:** Valor de distribución acumulativo de uno menos el porcentaje de hogares con cualquier deuda relacionada con el hogar. El valor se utiliza como característica de rendimiento.
- **hs_degree:** Porcentaje de personas con al menos un título de bachillerato.
- **hc_mortgage_mean:** Los costos hipotecarios mensuales medios y del propietario de una ubicación geográfica específica.
- **hc_mortgage_median:** la mediana de los costos hipotecarios mensuales y del propietario del área geográfica especificada.
- **hc_mortgage_stdev:** La desviación estándar de la hipoteca mensual y los costos del propietario para una ubicación geográfica especificada.
- **hc_mortgage_sample:** El número de muestras utilizadas en los cálculos estadísticos.
- **hc_mean:** Los costos mensuales medios del propietario de una ubicación geográfica especificada.
- **hc_median:** La mediana de los costos mensuales del propietario de una ubicación geográfica específica.
- **hc_stdev:** La desviación estándar de los costos mensuales del propietario de un área geográfica especificada.
- **hc_samples:** Las muestras utilizadas en el cálculo de las estadísticas de los costos mensuales del propietario.
- **rent_mean:** La renta bruta media de la ubicación geográfica especificada.
- **rent_stdev:** La desviación estándar del alquiler bruto para la ubicación geográfica especificada.

- **rent_samples:** El número de registros de alquiler bruto utilizados en los cálculos estadísticos.
- **rent_sample_weight:** La suma del peso bruto del alquiler utilizado en los cálculos.
- **rent_gt_10 (CDF):** El valor de distribución empírico de que el alquiler de un individuo será superior al 10% de sus ingresos familiares en los últimos 12 meses.
- **rent_gt_15 (CDF):** El valor de distribución empírico de que el alquiler de un individuo será superior al 15% de sus ingresos familiares en los últimos 12 meses.
- **rent_gt_20 (CDF):** El valor de distribución empírico de que el alquiler de una persona será superior al 20% de sus ingresos familiares en los últimos 12 meses.
- **rent_gt_25 (CDF):** El valor de distribución empírico de que el alquiler de un individuo será superior al 25% de sus ingresos familiares en los últimos 12 meses.
- **rent_gt_30 (CDF):** El valor de distribución empírico de que el alquiler de un individuo será superior al 30% de sus ingresos familiares en los últimos 12 meses.
- **rent_gt_35 (CDF):** El valor de distribución empírica de que el alquiler de un individuo será superior al 35% de sus ingresos familiares en los últimos 12 meses.
- **rent_gt_40 (CDF):** El valor de distribución empírica de que el alquiler de un individuo será superior al 40% de sus ingresos familiares en los últimos 12 meses.
- **rent_gt_50 (CDF):** El valor de distribución empírico de que el alquiler de una persona será superior al 50% de sus ingresos familiares en los últimos 12 meses.
- **rent_universe_samples (CDF):** El tamaño de las unidades de vivienda ocupadas por inquilinos muestreadas para el cálculo.
- **ingresos_familia_mean:** El ingreso familiar promedio de la ubicación geográfica especificada.
- **family_income_median:** El ingreso familiar promedio de la ubicación geográfica especificada.
- **family_income_stdev:** La desviación estándar del ingreso familiar para el área geográfica especificada.
- **family_income_families:** El número de familias utilizadas en los cálculos estadísticos.
- **hi_mean:** El ingreso familiar promedio de la ubicación geográfica especificada.
- **hi_stdev:** La desviación estándar del ingreso familiar para el área geográfica especificada.

Estadísticos Básicos.

Presentamos un resumen de cada columna, esta nos muestra las estadísticas básicas que podrían otorgarnos información, pero debemos tener cuidado pues algunos estadísticos pueden estar erróneos por los valores atípicos que suelen aparecer en las bases de datos.

UID	BLOCKID	SUMLEVEL	COUNTYID	STATEID
Min. :220336	Mode:logical	Min. :140	Min. : 1.00	Min. :1.000
1st Qu.:222824	NA's:5292	1st Qu.:140	1st Qu.: 21.00	1st Qu.:4.000
Median :225394		Median :140	Median : 37.00	Median :6.000
Mean :225387		Mean :140	Mean : 48.47	Mean :4.972
3rd Qu.:227963		3rd Qu.:140	3rd Qu.: 71.00	3rd Qu.:6.000
Max. :230479		Max. :140	Max. :290.00	Max. :6.000

state	state_ab	city	place
Length:5292	Length:5292	Length:5292	Length:5292
Class :character	Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character	Mode :character

type	primary	zip_code	area_code
Length:5292	Length:5292	Min. :32445	Min. :205
Class :character	Class :character	1st Qu.:85283	1st Qu.:415
Mode :character	Mode :character	Median :91342	Median :602
		Mean :83677	Mean :603
		3rd Qu.:92869	3rd Qu.:818
		Max. :99929	Max. :951

lat	lng	ALand	AWater
Min. :30.24	Min. :-166.77	Min. :6.128e+04	Min. :0.000e+00
1st Qu.:33.54	1st Qu.: -118.50	1st Qu.:1.284e+06	1st Qu.:0.000e+00
Median :34.03	Median : -117.75	Median :2.633e+06	Median :0.000e+00
Mean :35.13	Mean : -112.95	Mean :2.305e+08	Mean :2.734e+07
3rd Qu.:35.39	3rd Qu.: -111.85	3rd Qu.:1.353e+07	3rd Qu.:8.378e+04
Max. :67.07	Max. : -84.98	Max. :1.040e+11	Max. :2.453e+10

pop	male_pop	female_pop	rent_mean
Min. : 0	Min. : 0	Min. : 0	Min. : 181.8
1st Qu.: 3240	1st Qu.: 1582	1st Qu.: 1620	1st Qu.: 879.3
Median : 4345	Median : 2137	Median : 2200	Median :1177.0
Mean : 4603	Mean : 2293	Mean : 2310	Mean :1257.1
3rd Qu.: 5652	3rd Qu.: 2804	3rd Qu.: 2870	3rd Qu.:1535.1
Max. :39454	Max. :27962	Max. :12537	Max. :3962.3

rent_median	rent_stdev	rent_sample_weight	rent_samples
Min. : 131	Min. : 18.26	Min. : 0.3928	Min. : 4.0
1st Qu.: 828	1st Qu.: 312.00	1st Qu.: 90.7074	1st Qu.: 280.0
Median :1116	Median : 416.52	Median : 204.5361	Median : 529.0
Mean :1209	Mean : 454.64	Mean : 276.3538	Mean : 642.5
3rd Qu.:1474	3rd Qu.: 557.01	3rd Qu.: 375.8788	3rd Qu.: 875.0
Max. :3972	Max. :1556.38	Max. :3060.2479	Max. :6281.0
NA's :55	NA's :55	NA's :55	NA's :55

rent_gt_10	rent_gt_15	rent_gt_20	rent_gt_25
Min. :0.3042	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.9538	1st Qu.:0.8501	1st Qu.:0.7074	1st Qu.:0.5631
Median :0.9838	Median :0.9109	Median :0.7958	Median :0.6684
Mean :0.9657	Mean :0.8897	Mean :0.7737	Mean :0.6516
3rd Qu.:1.0000	3rd Qu.:0.9547	3rd Qu.:0.8661	3rd Qu.:0.7568
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
NA's :55	NA's :55	NA's :55	NA's :55

rent_gt_30	rent_gt_35	rent_gt_40	rent_gt_50
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.4415	1st Qu.:0.3455	1st Qu.:0.2752	1st Qu.:0.1854
Median :0.5487	Median :0.4539	Median :0.3761	Median :0.2718
Mean :0.5381	Mean :0.4464	Mean :0.3749	Mean :0.2757
3rd Qu.:0.6458	3rd Qu.:0.5509	3rd Qu.:0.4756	3rd Qu.:0.3580
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
NA's :55	NA's :55	NA's :55	NA's :55

universe_samples	used_samples	hi_mean	hi_median
Min. : 0.0	Min. : 0.0	Min. : 5000	Min. : 4790
1st Qu.: 303.0	1st Qu.: 265.0	1st Qu.: 49494	1st Qu.: 37137
Median : 559.5	Median : 515.0	Median : 66292	Median : 52674
Mean : 666.8	Mean : 620.6	Mean : 73319	Mean : 59944
3rd Qu.: 899.0	3rd Qu.: 857.0	3rd Qu.: 92037	3rd Qu.: 76596
Max. : 6648.0	Max. : 6094.0	Max. : 204239	Max. : 217166
		NA's : 46	NA's : 46
hi_stdev	hi_sample_weight	hi_samples	family_mean
Min. : 2451	Min. : 1.2	Min. : 8	Min. : 19353
1st Qu.: 42845	1st Qu.: 587.6	1st Qu.: 1107	1st Qu.: 54959
Median : 53988	Median : 835.4	Median : 1491	Median : 72504
Mean : 56412	Mean : 899.2	Mean : 1588	Mean : 79795
3rd Qu.: 68390	3rd Qu.: 1140.7	3rd Qu.: 1963	3rd Qu.: 100051
Max. : 135903	Max. : 4991.1	Max. : 7177	Max. : 199750
NA's : 46	NA's : 46	NA's : 46	NA's : 47
family_median	family_stdev	family_sample_weight	family_samples
Min. : 9051	Min. : 1826	Min. : 1.2	Min. : 4
1st Qu.: 43587	1st Qu.: 40972	1st Qu.: 345.1	1st Qu.: 724
Median : 60900	Median : 50598	Median : 510.7	Median : 1003
Mean : 69741	Mean : 51848	Mean : 550.1	Mean : 1076
3rd Qu.: 88427	3rd Qu.: 62697	3rd Qu.: 699.0	3rd Qu.: 1339
Max. : 220353	Max. : 111257	Max. : 4888.9	Max. : 6603
NA's : 47	NA's : 47	NA's : 47	NA's : 47
hc_mortgage_mean	hc_mortgage_median	hc_mortgage_stdev	
Min. : 526.3	Min. : 370	Min. : 36.51	
1st Qu.: 1287.4	1st Qu.: 1185	1st Qu.: 506.63	
Median : 1802.3	Median : 1710	Median : 695.95	
Mean : 1871.1	Mean : 1806	Mean : 700.58	
3rd Qu.: 2328.2	3rd Qu.: 2261	3rd Qu.: 889.83	
Max. : 4462.3	Max. : 4472	Max. : 1501.78	
NA's : 108	NA's : 108	NA's : 108	
hc_mortgage_sample_weight	hc_mortgage_samples	hc_mean	hc_median
Min. : 0.1984	Min. : 1.0	Min. : 53.59	Min. : 53
1st Qu.: 114.2055	1st Qu.: 321.0	1st Qu.: 371.28	1st Qu.: 340
Median : 197.1278	Median : 551.0	Median : 456.72	Median : 421
Mean : 237.6411	Mean : 630.3	Mean : 501.65	Mean : 470
3rd Qu.: 312.3094	3rd Qu.: 843.0	3rd Qu.: 587.07	3rd Qu.: 550
Max. : 1493.8975	Max. : 4000.0	Max. : 1549.56	Max. : 1674
NA's : 108	NA's : 108	NA's : 127	NA's : 127
hc_stdev	hc_samples	hc_sample_weight	
Min. : 18.26	Min. : 2.0	Min. : 0.9899	
1st Qu.: 148.46	1st Qu.: 130.0	1st Qu.: 87.5293	
Median : 198.82	Median : 234.0	Median : 159.2020	
Mean : 219.68	Mean : 297.2	Mean : 213.4245	
3rd Qu.: 271.79	3rd Qu.: 393.0	3rd Qu.: 275.8990	
Max. : 820.97	Max. : 2807.0	Max. : 2088.2890	
NA's : 127	NA's : 127	NA's : 127	
home_equity_second_mortgage	second_mortgage	home_equity	
Min. : 0.00000	Min. : 0.00000	Min. : 0.00000	
1st Qu.: 0.00637	1st Qu.: 0.00965	1st Qu.: 0.05187	
Median : 0.02259	Median : 0.02780	Median : 0.10203	
Mean : 0.03032	Mean : 0.03608	Mean : 0.10909	
3rd Qu.: 0.04408	3rd Qu.: 0.05214	3rd Qu.: 0.15620	
Max. : 0.43750	Max. : 0.56000	Max. : 1.00000	
NA's : 80	NA's : 80	NA's : 80	
debt	second_mortgage_cdf	home_equity_cdf	debt_cdf
Min. : 0.0000	Min. : 0.00053	Min. : 0.0000	Min. : 0.0000
1st Qu.: 0.5858	1st Qu.: 0.19405	1st Qu.: 0.2242	1st Qu.: 0.1749
Median : 0.7055	Median : 0.36821	Median : 0.4325	Median : 0.3550
Mean : 0.6748	Mean : 0.42978	Mean : 0.4525	Mean : 0.4108
3rd Qu.: 0.7898	3rd Qu.: 0.54002	3rd Qu.: 0.6666	3rd Qu.: 0.6291
Max. : 1.0000	Max. : 1.00000	Max. : 1.0000	Max. : 1.0000
NA's : 80	NA's : 80	NA's : 80	NA's : 80

hs_degree	hs_degree_male	hs_degree_female	male_age_mean
Min. :0.2369	Min. :0.0000	Min. :0.1997	Min. :18.65
1st Qu.:0.7414	1st Qu.:0.7311	1st Qu.:0.7465	1st Qu.:33.59
Median :0.8578	Median :0.8570	Median :0.8621	Median :37.19
Mean :0.8174	Mean :0.8149	Mean :0.8206	Mean :37.73
3rd Qu.:0.9345	3rd Qu.:0.9380	3rd Qu.:0.9344	3rd Qu.:40.84
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :77.76
NA's :28	NA's :29	NA's :32	NA's :27
male_age_median	male_age_stdev	male_age_sample_weight	male_age_samples
Min. :15.17	Min. :0.7371	Min. :2.87	Min. :6
1st Qu.:31.00	1st Qu.:19.9726	1st Qu.:391.77	1st Qu.:1590
Median :36.00	Median :21.3609	Median :530.02	Median :2141
Mean :37.01	Mean :21.0667	Mean :581.28	Mean :2305
3rd Qu.:41.67	3rd Qu.:22.6037	3rd Qu.:706.46	3rd Qu.:2806
Max. :78.92	Max. :27.6785	Max. :12017.07	Max. :27962
NA's :27	NA's :27	NA's :27	NA's :27
female_age_mean	female_age_median	female_age_stdev	female_age_sample_weight
Min. :19.44	Min. :15.33	Min. :1.466	Min. :0.746
1st Qu.:35.33	1st Qu.:33.08	1st Qu.:20.804	1st Qu.:393.258
Median :39.09	Median :38.33	Median :22.030	Median :538.848
Mean :39.53	Mean :39.11	Mean :21.734	Mean :573.405
3rd Qu.:43.01	3rd Qu.:44.33	3rd Qu.:23.149	3rd Qu.:710.424
Max. :77.00	Max. :78.00	Max. :28.180	Max. :3406.539
NA's :30	NA's :30	NA's :30	NA's :30
female_age_samples	pct_own	married	married_snp
Min. :4	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:1628	1st Qu.:0.4172	1st Qu.:0.4175	1st Qu.:0.02705
Median :2207	Median :0.6117	Median :0.5023	Median :0.04753
Mean :2324	Mean :0.5750	Mean :0.4964	Mean :0.05466
3rd Qu.:2875	3rd Qu.:0.7593	3rd Qu.:0.5831	3rd Qu.:0.07497
Max. :12537	Max. :1.0000	Max. :1.0000	Max. :0.71429
NA's :30	NA's :46	NA's :28	NA's :28

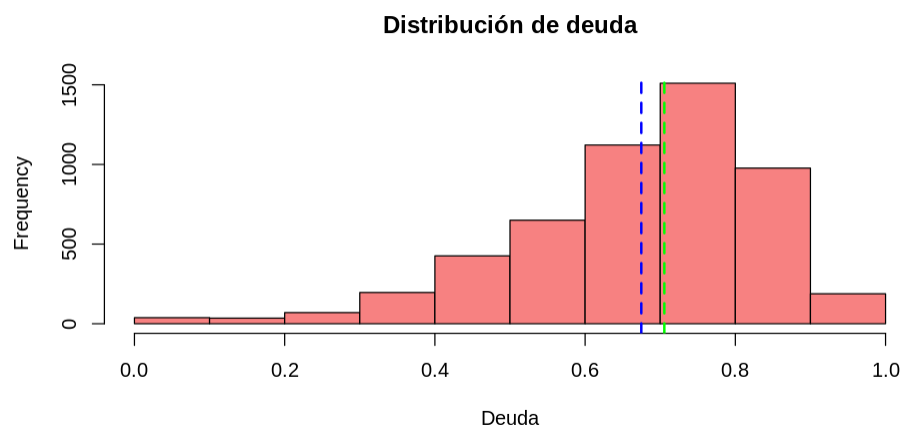
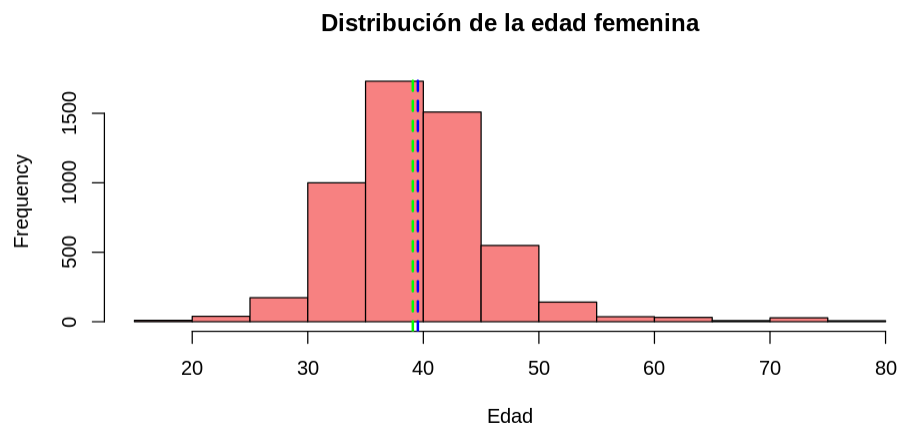
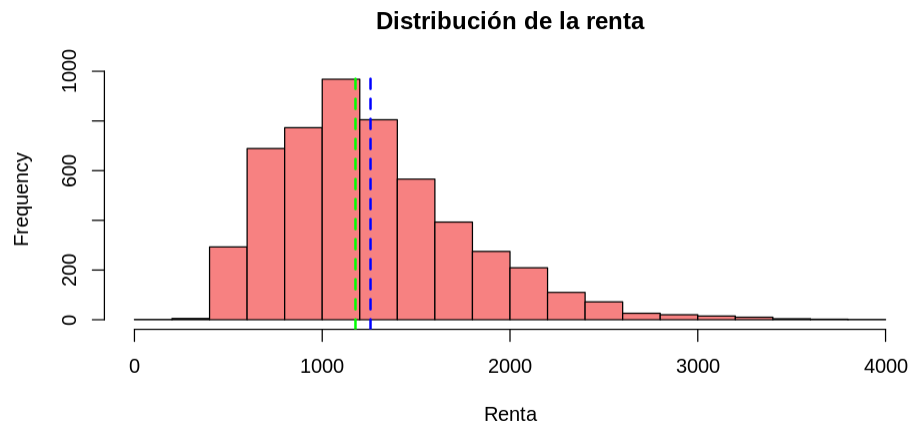
separated	divorced
Min. :0.000000	Min. :0.00000
1st Qu.:0.005795	1st Qu.:0.05891
Median :0.015145	Median :0.08635
Mean :0.019962	Mean :0.09330
3rd Qu.:0.028285	3rd Qu.:0.12159
Max. :0.714290	Max. :0.33565
NA's :28	NA's :28

Tipos de Distribución:

Tomaremos variables de nuestra base de datos para poder explicar cada una de las distribuciones, podemos observar que contamos con los tres tipos de distribución principal (normal, sesgo hacia la derecha e izquierda).

- **Distribución normal:** También conocida como curva de campana, es una distribución en la que la mitad de los datos se encuentra en el lado izquierdo y la otra mitad se encuentra en el lado derecho de la distribución. En esta distribución, la curva es simétrica y la media, la moda y la mediana son todas iguales.
- **Distribución sesgada a la derecha:** tiene una cola larga que apunta a la derecha. Esto significa que, en nuestra muestra o población, la mayoría de los datos se concentran en el lado izquierdo de la distribución. La media de la renta está en este caso sesgada a la derecha, lo que nos indica que la renta media se concentró principalmente en el lado izquierdo, lo que significa que la mayoría de las observaciones no pueden permitirse una renta alta.
- **Distribución sesgada a la izquierda:** tiene una cola larga que apunta a la izquierda. Esto significa que, en nuestra muestra o población, la mayoría de los datos se concentran en el lado derecho de la distribución. Deuda es un ejemplo de una distribución sesgada a la izquierda, lo que significa que la

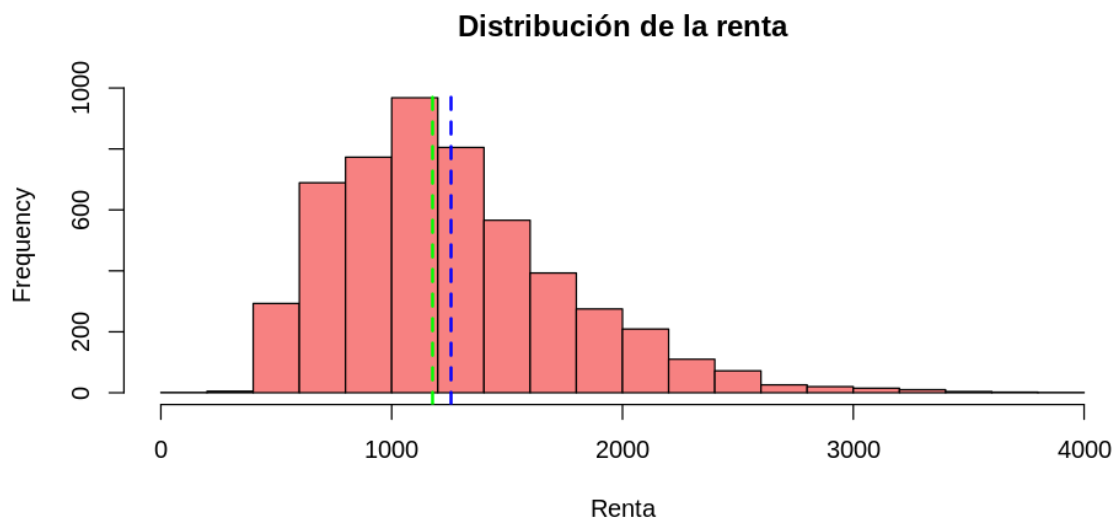
mayoría de las observaciones tienen una alta concentración de deuda, lo que significa que la mayoría de las observaciones están en el lado derecho.



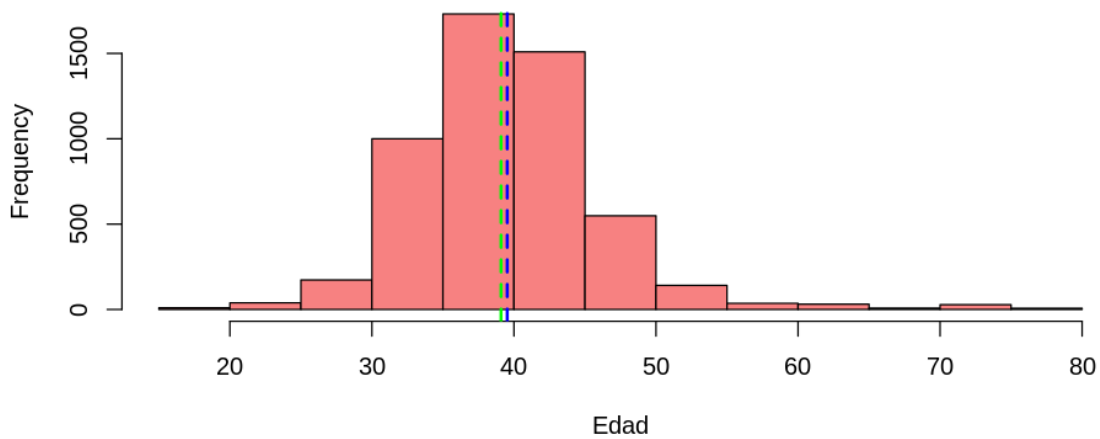
Mostramos un gráfico para demostrar el sesgo de nuestras distribuciones, sabemos que, si la mediana se encuentra en el mismo lugar que la media, contamos con una distribución normal.

	debt	rent_mean	female_age_mean
Min.	0.0000000	181.7723	19.76781
1st Qu.	0.5860650	878.9692	35.40263
Median	0.7054000	1177.0329	39.13049
Mean	0.6749746	1257.1211	39.58056
3rd Qu.	0.7897550	1534.8831	43.02770
Max.	1.0000000	3962.3423	76.99683

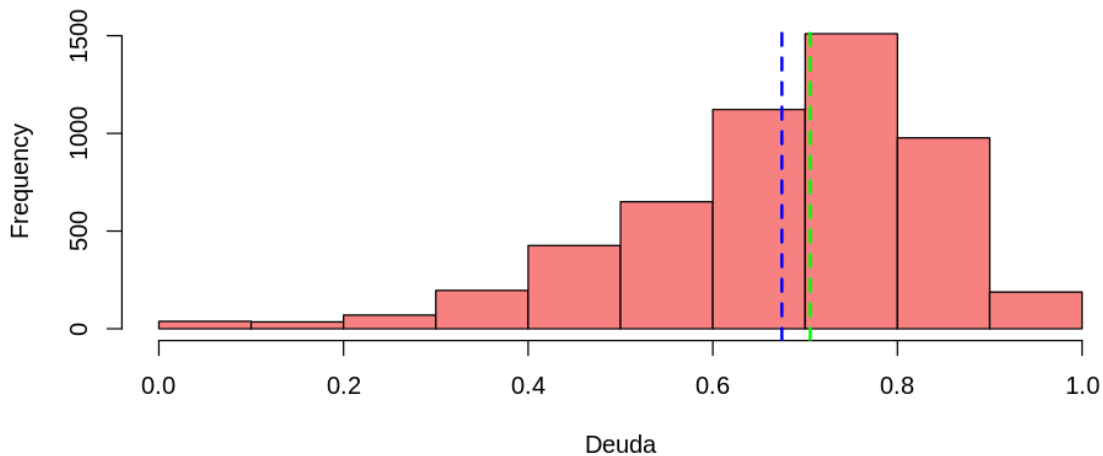
Mostramos un gráfico similar al anterior para demostrar el sesgo de nuestras distribuciones, sabemos que, si la mediana se encuentra en el mismo lugar que la media, contamos con una distribución normal.



Distribución de la edad femenina



Distribución de deuda



Otras medidas estadísticas:

- **Varianza:** Este es un indicador de cómo se distribuyen nuestros datos. La variabilidad más pequeña que podría haber es 0 mientras que la mayor es infinita. La varianza se expresa como:

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N}$$

- **Desviación estándar:** La desviación estándar es solo la raíz cuadrada de nuestra varianza y nos dice qué tan lejos están nuestros datos de la media.

La desviación estándar se expresa como:

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

- **1er Cuartil(Q1):** Esto está compuesto por el 25% más bajo de números en nuestra distribución.
- **2do Cuartil (Q2):** Compuesto por el 50% de los números más bajos hasta la **mediana**.
- **3rd Cuartil (Q3):** Compuesto por el 75% de los números más bajos.
- **Rango intercuartil (IQR):** Esto nos ayuda a detectar dónde se encuentran la mayoría de los datos. IQR se expresa como:
IQR = Q1-Q3
Se prefiere utilizar IQR en lugar de la media o la mediana cuando se trata de averiguar dónde se encuentran la mayoría de los datos.

Ahora haremos una tabla que nos muestre cuartiles y rango intercuartil de cada variable.

```
[1] "Estadísticas de renta media"
[1] "-----"

|      mu| rent_med|      std| rent_min| rent_max|  rent_q1|  rent_q3| rent_iqr|
|-----:|-----:|-----:|-----:|-----:|-----:|-----:|-----:|
| 1257.125| 1177.033| 505.5629| 181.7723| 3962.342| 879.2731| 1535.102| 655.8291|

[1] "Estadísticas medias de edad femenina"
[1] "-----"

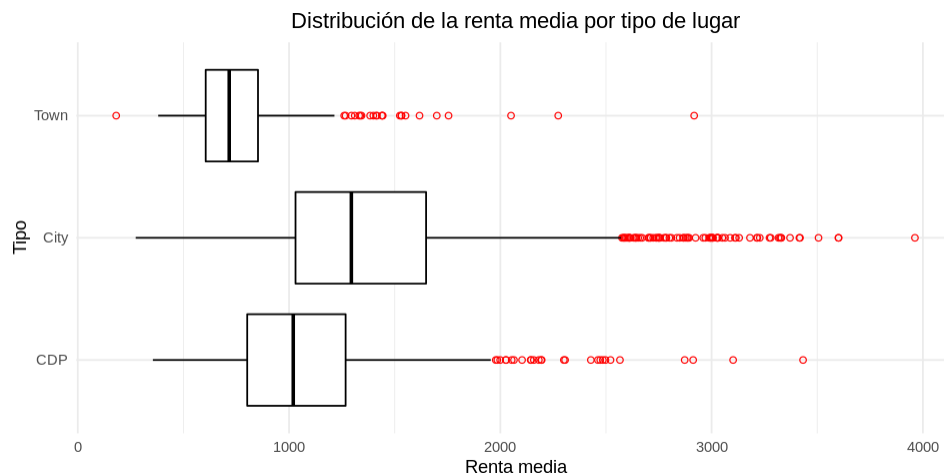
|      mu| fage_med|      std| fage_min| fage_max| fage_q1| fage_q3| fage_iqr|
|-----:|-----:|-----:|-----:|-----:|-----:|-----:|-----:|
| 39.52573| 39.08989| 6.635169| 19.4442| 76.99683| 35.3303| 43.00736| 7.67706|

[1] "Estadísticas de deuda"
[1] "-----"

|      mu| debt_med|      std| debt_min| debt_max|  debt_q1|  debt_q3| debt_iqr|
|-----:|-----:|-----:|-----:|-----:|-----:|-----:|-----:|
| 0.6748094| 0.70547| 0.1646559| 0| 1| 0.585815| 0.789845| 0.20403|
```

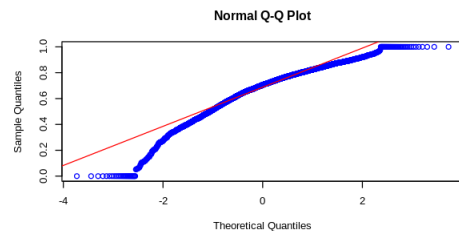
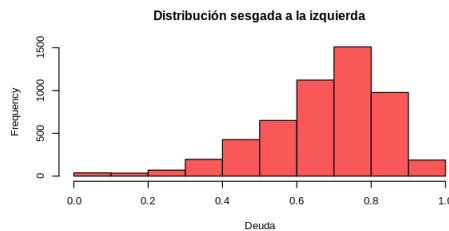
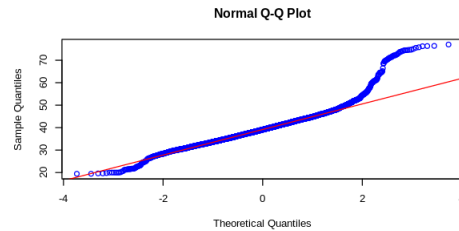
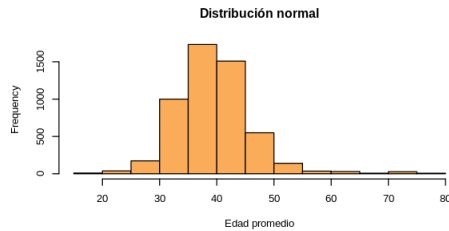
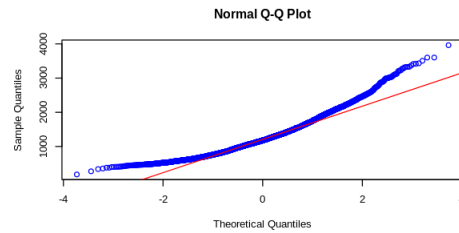
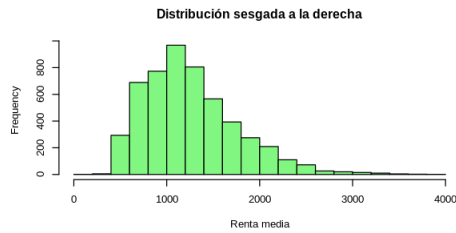
Diagramas de caja y presuntos valores atípicos.

Solo queríamos agregar en esta sección que los valores atípicos deben analizarse cuidadosamente y, aunque existen reglas comunes como la de una "distribución normal", cualquier valor más allá de tres desviaciones estándar debe considerarse un valor atípico. Aunque existe una pequeña probabilidad de que un valor en una distribución normal esté a 3 desviaciones estándar de la media, debemos analizar cuidadosamente por qué es así. Podría ser que los datos estuvieran mal escritos, lo que debilitaría la teoría de que una observación específica es un valor atípico.



Q-Q Plots.

- **Gráfica Q sesgada a la derecha:** Cuando la distribución está sesgada a la derecha, las observaciones tienden a ir por encima de la línea roja, lo que indica que la distribución está sesgada a la derecha.
- **Q-plot de distribución normal:** Aunque algunas observaciones no están en la línea, la mayoría de las observaciones están en la línea, lo que indica que la distribución es mayormente normal.
- **Gráfica Q sesgada a la izquierda:** Aunque esta distribución no está fuertemente sesgada a la izquierda, podemos ver que la mayoría de las observaciones caen por debajo de la línea roja, lo que indica que la distribución está sesgada a la izquierda.

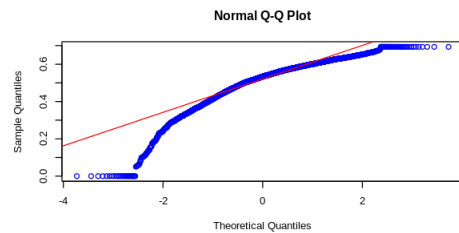
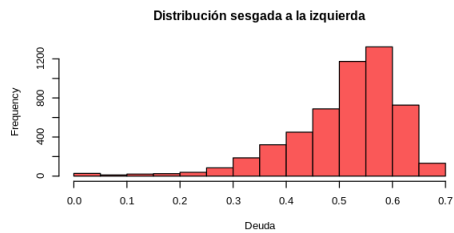
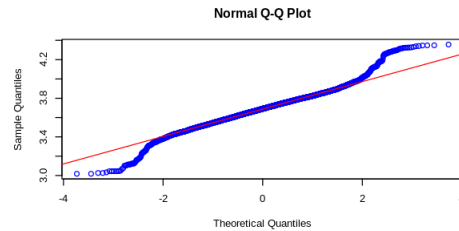
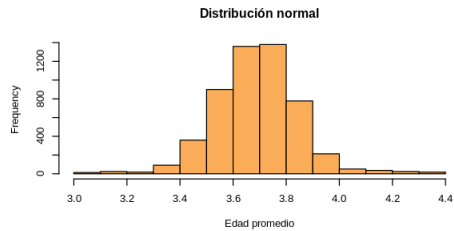
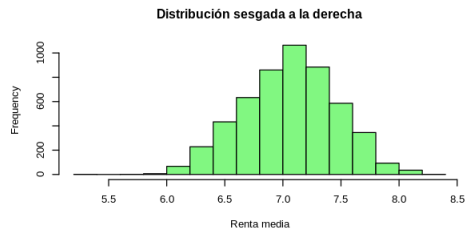


Usar logaritmos naturales para crear distribuciones normales.

Aunque a veces el uso de logaritmos naturales no necesariamente impacta en las observaciones para formar una distribución normal (ejemplo sesgado a la izquierda), la mayoría de las veces nos da una distribución normal aproximada como en el ejemplo sesgado a la derecha.

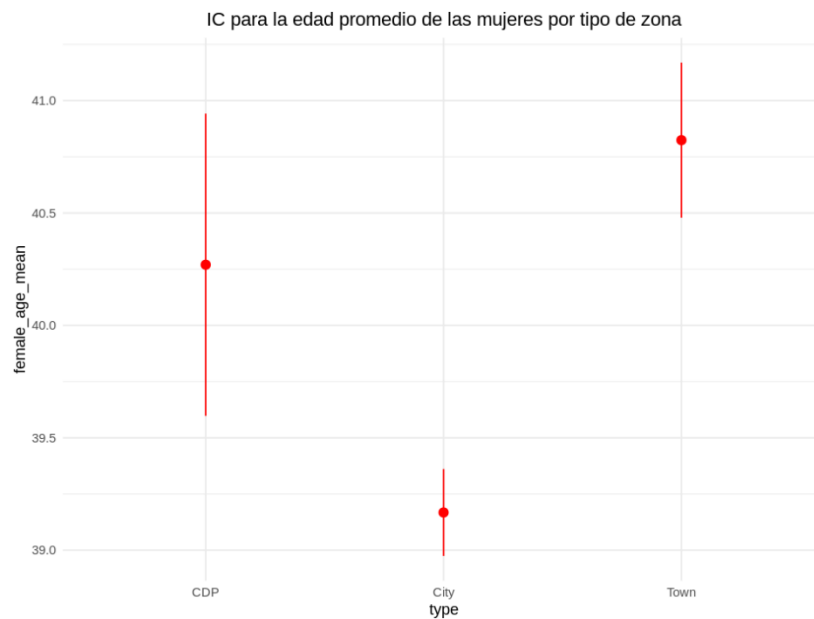
¿Por qué querríamos una distribución normal?

Aunque hay muchas formas de lidiar con la asimetría, la mayoría de las técnicas estadísticas asumen que la distribución es "normal". Explicaremos con más detalle qué técnicas estadísticas son esas. Las pruebas estadísticas como las pruebas z, t y F suponen que la media tiene una distribución "normal". Además, es algo más sencillo calcular probabilidades y calcular intervalos de confianza asumiendo que la distribución cumple las condiciones del Teorema del límite central (TLC).



Visualización de intervalos de confianza con ggplot:

Con los intervalos de confianza, nos aseguramos de nuestra confianza en cuál es el promedio real de la población. **Cuanto más anchas sean las barras de error, menos seguros de cuál es la media real.**



ymin	ymax
<dbl>	<dbl>
19.4442	76.99683

	median	first_quartile	third_quartile	interquartile_range
	<dbl>	<dbl>	<dbl>	<dbl>
25%	39.08988	35.3303	43.00736	7.67706

Estadísticas de inferencia:

En esta sección hablaremos sobre la importancia de los intervalos de confianza y cómo podemos encontrar el intervalo de confianza de una distribución. Además, entraremos brevemente en los temas de la prueba de hipótesis y explicaremos por qué es importante conocer estos conceptos antes incluso de intentar responder a la pregunta de nuestro problema principal.

Prueba de hipótesis.

Imagine un escenario en el que un individuo está en un juicio por cometer un asesinato en los Estados Unidos. Hasta donde sabemos, cuando un individuo es considerado "inocente" hasta que se demuestre lo contrario. A través de este breve ejemplo, me gustaría presentar el concepto de Prueba de hipótesis. Hay dos tipos de hipótesis: la hipótesis nula y la hipótesis alternativa.

Existen dos tipos de hipótesis: la hipótesis nula y la hipótesis alternativa.

- **Hipótesis nula (H0):** Este es el "estatus quo". Cuando queremos comparar medias de dos variables, digamos el ingreso promedio de hombres y mujeres, la hipótesis Nula en este caso será que "no hay diferencia".
- **Hipótesis alternativa (HA):** Esto va en contra de lo que afirmaba la Hipótesis nula. En nuestro ejemplo de ingresos promedio por género, el ingreso promedio de los hombres no es igual al ingreso promedio de las mujeres.

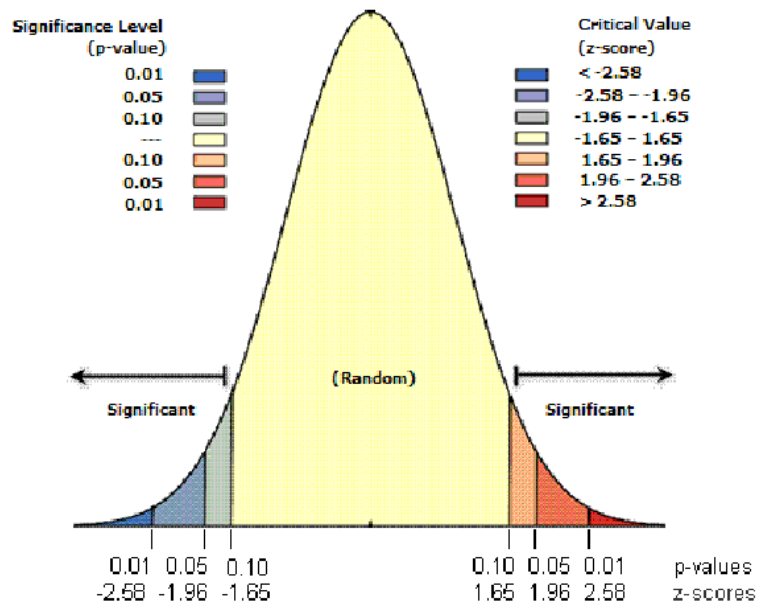
Hipótesis nula (ejemplo de ingresos medios):

$$A_{income_M} = A_{income_F}$$

Hipótesis alternativa (ejemplo de ingreso promedio):

$$A_{income_M} \neq A_{income_F}$$

Intervalos de confianza y valores p:



Los intervalos de confianza (IC): son la certeza de que un valor específico se ubicará entre dos puntos específicos. Los tipos de intervalos de confianza más habituales son los intervalos de confianza del 90%, 95% y 99% aunque el 95% es el que más se utiliza y es el que usaremos en este ejemplo.

Valor p: es la probabilidad de que ocurra un evento dado. Suponiendo que un intervalo de confianza es del 95%, si el valor p, rechazamos la Hipótesis nula a favor de la Hipótesis alternativa.

Nivel de significancia: es la probabilidad de rechazar la Hipótesis nula (también denotada como alfa o α).

Hallar el intervalo de confianza para la media poblacional:

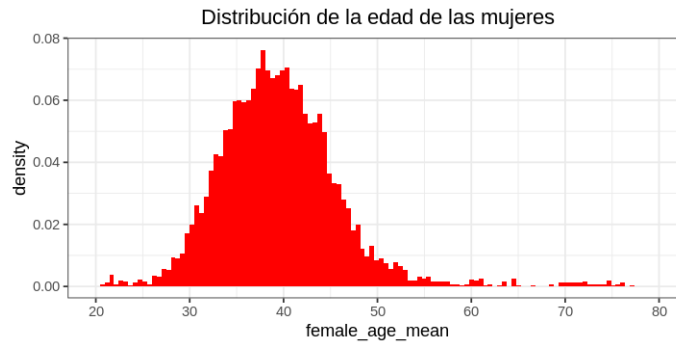
$$\bar{x} \pm z^* s \frac{s}{\sqrt{N-1}}$$

\bar{x} = Muestra promedio

z = puntuación z

s = Error estándar

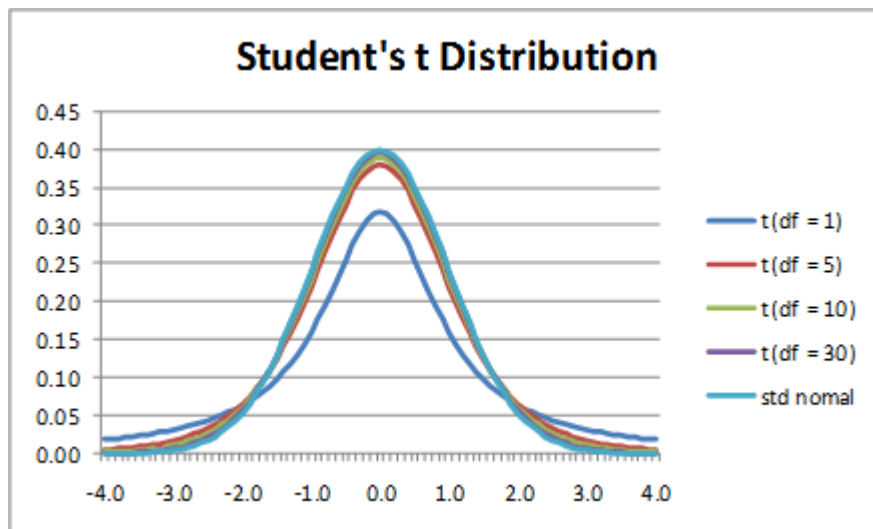
N = Tamaño de la muestra



- Número de filas después de filtrar valores nulos 5262
- El intervalo inferior es 39.35.
- El intervalo superior es: 39.71
- El porcentaje en el que el promedio real es intermedio: 90%
- Z_score es: -0.24
- El valor p de dos colas es: 0.81
- El valor p unilateral es: 0.41

En este caso, dado que el p-valor $> \alpha$ no rechazamos H_0 .

Distribución T



La distribución t es una distribución que solo se usa para muestras pequeñas. La primera pregunta que tuve al tratar con este tipo de distribución es por qué necesitamos una distribución t cuando recibimos toneladas diarias de datos, lo que hace imposible tener una muestra pequeña. Bueno, las distribuciones t se usan con más frecuencia cuando se realiza un experimento que suele tener muestras más pequeñas.

- **Tamaño de la muestra:** El tamaño de la muestra debe ser menor que 30 para ser considerado para una distribución t.

- **Grados de libertad:** A medida que el tamaño de la muestra se acerca a 30, la distribución t se verá exactamente como una distribución normal. Además, los grados de libertad determinan el grosor de la cola.
- **Puntuación T:** Para calcular la puntuación t usamos la siguiente fórmula, $(\bar{x} - \mu) / s$, donde s es el error estándar y μ la hipótesis nula.

Ejercicio 1: Hallemos el intervalo de confianza del 95% de una muestra para alquiler en el estado de Nueva York.

Tomamos los grados de libertad de nuestra n-1, nos resulta una t= 0.0025 de nuestra tabla t student.

$$\bar{x} \pm (0.0025) * (s / \sqrt{n})$$

Obteniendo el promedio de alquileres del estado de Nueva York, un dicho tamaño de muestra obtenemos con un 95% de confianza en que la renta media de la ciudad de Nueva York se encuentra entre 934 y 1417.

Ejercicio 2: Supongamos que $\mu = 1000$ para el alquiler de Nueva York, encontremos el valor p para ver si hay suficiente evidencia de que podríamos rechazar H_0 a favor de H_A , recuerde si el valor p es menor que 0.05 nivel de significancia luego rechazamos el Nulo a favor de la alternativa.

Con una $t = [(\bar{x} - 1000) / (s / \sqrt{n})] = 1.605$ con 11 grados de libertad.

$$\text{Valor P} = 2 * [1 - P(Z < |Z_{\text{calculada}}|)]$$

$$P_value = 0.136796175516726$$

Dado que nuestro valor p es mayor que 0.05, nos inclinamos a favor de la Hipótesis nula, lo que significa que no hay evidencia suficiente de que el alquiler promedio en Nueva York sea algo diferente a 1000.

CONCLUSION

La estadística inferencial es una rama de la estadística que estudia el comportamiento y propiedades de una muestra para poder generalizar unos resultados obtenidos, con esta magnífica rama de la estadística pudimos crear muestras aleatorias simples para poder determinar hipótesis que teníamos acerca de la base de datos. Este es un gran comienzo para el análisis de datos

Creemos que esta sección puede ser menos precisa a la hora de inferir o predecir, pero puede ser de gran ayuda como dato a la hora de toma de decisiones rápidas pero inteligente.

Podemos llevarnos como tarea en un futuro crear mejores técnicas de muestreo para evitar sesgos de precisión que pudieran provocar falta de precisión a la hora de modelos de predicción.

Teniendo en cuenta el punto anterior podemos aplicar más técnicas de limpieza de datos para mejorar cada inferencia para no perder mucha información pues vimos que en las fórmulas para realizar los intervalos de confianza y muestreo los datos NA simplemente fueron eliminados y no fueron transformados, como vimos en la gráfica estos mismos datos están normalmente distribuidos, pero podrían darnos información errónea el hecho de simplemente eliminarlos.

Anexos.

- Tabla de Frecuencia de los Estados.

Estado	Absoluta	Absoluta Acumulada	Relativa	Relativa Acumulada	Relativa Porcentual	Relativa Acumulada Porcentual
Alabama	612	612	0.01568	0.01568	1.57%	1.57%
Alaska	105	717	0.00269	0.01837	0.27%	1.84%
Arizona	798	1515	0.02045	0.03882	2.04%	3.88%
Arkansas	363	1878	0.00930	0.04812	0.93%	4.81%
California	4193	6071	0.10743	0.15555	10.74%	15.55%
Colorado	668	6739	0.01712	0.17266	1.71%	17.27%
Connecticut	445	7184	0.01140	0.18406	1.14%	18.41%
Delaware	109	7293	0.00279	0.18686	0.28%	18.69%
District of Columbia	98	7391	0.00251	0.18937	0.25%	18.94%
Florida	2289	9680	0.05865	0.24801	5.86%	24.80%
Georgia	1078	10758	0.02762	0.27563	2.76%	27.56%
Hawaii	174	10932	0.00446	0.28009	0.45%	28.01%
Idaho	148	11080	0.00379	0.28388	0.38%	28.39%
Illinois	1593	12673	0.04081	0.32470	4.08%	32.47%
Indiana	802	13475	0.02055	0.34525	2.05%	34.52%
Iowa	415	13890	0.01063	0.35588	1.06%	35.59%
Kansas	440	14330	0.01127	0.36715	1.13%	36.72%
Kentucky	577	14907	0.01478	0.38194	1.48%	38.19%
Louisiana	608	15515	0.01558	0.39751	1.56%	39.75%
Maine	190	15705	0.00487	0.40238	0.49%	40.24%
Maryland	764	16469	0.01957	0.42196	1.96%	42.20%
Massachusetts	777	17246	0.01991	0.44187	1.99%	44.19%
Michigan	1463	18709	0.03748	0.47935	3.75%	47.93%
Minnesota	699	19408	0.01791	0.49726	1.79%	49.73%
Mississippi	351	19759	0.00899	0.50625	0.90%	50.63%
Missouri	732	20491	0.01875	0.52501	1.88%	52.50%
Montana	161	20652	0.00413	0.52913	0.41%	52.91%
Nebraska	274	20926	0.00702	0.53615	0.70%	53.62%
Nevada	355	21281	0.00910	0.54525	0.91%	54.52%
New Hampshire	157	21438	0.00402	0.54927	0.40%	54.93%
New Jersey	1009	22447	0.02585	0.57512	2.59%	57.51%
New Mexico	287	22734	0.00735	0.58248	0.74%	58.25%
New York	2565	25299	0.06572	0.64819	6.57%	64.82%
North Carolina	1167	26466	0.02990	0.67809	2.99%	67.81%
North Dakota	110	26576	0.00282	0.68091	0.28%	68.09%
Ohio	1538	28114	0.03941	0.72032	3.94%	72.03%
Oklahoma	552	28666	0.01414	0.73446	1.41%	73.45%
Oregon	454	29120	0.01163	0.74609	1.16%	74.61%
Pennsylvania	1735	30855	0.04445	0.79055	4.45%	79.05%
Puerto Rico	489	31344	0.01253	0.80307	1.25%	80.31%
Rhode Island	128	31472	0.00328	0.80635	0.33%	80.64%
South Carolina	561	32033	0.01437	0.82073	1.44%	82.07%
South Dakota	128	32161	0.00328	0.82401	0.33%	82.40%
Tennessee	792	32953	0.02029	0.84430	2.03%	84.43%
Texas	2767	35720	0.07089	0.91519	7.09%	91.52%
Utah	331	36051	0.00848	0.92367	0.85%	92.37%
Vermont	95	36146	0.00243	0.92611	0.24%	92.61%
Virginia	1006	37152	0.02578	0.95188	2.58%	95.19%
Washington	806	37958	0.02065	0.97253	2.07%	97.25%
West Virginia	265	38223	0.00679	0.97932	0.68%	97.93%
Wisconsin	732	38955	0.01875	0.99808	1.88%	99.81%
Wyoming	75	39030	0.00192	1.00000	0.19%	100.00%

- Tabla de Frecuencia de la Renta Promedio (renta_mean).

n=	38,538	*El total de datos son 39,030, pero se quitaron 492 datos atípicos
k=	196.3109778	
Dato Min	117.15	
Dato Max	3962.34229	
Ancho de clase	19.61832801	

Intervalo	Marca de clase (xi)	Absoluta (ni)	Absoluta Acumulada	Relativa	Relativa Acumulada	Relativa Porcentual	Relativa acumulada Porcentual
117.15	136.768328	126.959164	2	5.18968E-05	5.18968E-05	0.005%	0.005%
136.768328	156.386656	146.577492	2	5.18968E-05	0.000103794	0.005%	0.010%
156.386656	176.004984	166.19582	3	7.78452E-05	0.000181639	0.008%	0.018%
176.004984	195.623312	185.814148	3	7.78452E-05	0.000259484	0.008%	0.026%
195.623312	215.2416401	205.432476	1	2.59484E-05	0.000285433	0.003%	0.029%
215.2416401	234.8599681	225.050804	7	0.000181639	0.000467071	0.018%	0.047%
234.8599681	254.4782961	244.669132	10	0.000259484	0.000726556	0.026%	0.073%
254.4782961	274.0966241	264.28746	8	0.000207587	0.000934143	0.021%	0.093%
274.0966241	293.7149521	283.905788	14	0.000363278	0.001297421	0.036%	0.130%
293.7149521	313.3332801	303.524116	8	0.000207587	0.001505008	0.021%	0.151%
313.3332801	332.9516081	323.142444	15	0.000389226	0.001894234	0.039%	0.189%
332.9516081	352.5699361	342.760772	23	0.000596814	0.002491048	0.060%	0.249%
352.5699361	372.1882641	362.3791	40	0.001037937	0.003528984	0.104%	0.353%
372.1882641	391.8065921	381.997428	46	0.001193627	0.004722611	0.119%	0.472%
391.8065921	411.4249202	401.615756	56	0.001453111	0.006175723	0.145%	0.618%
411.4249202	431.0432482	421.234084	73	0.001894234	0.008069957	0.189%	0.807%
431.0432482	450.6615762	440.852412	95	0.002465099	0.010535056	0.247%	1.054%
450.6615762	470.2799042	460.47074	145	0.00376252	0.014297576	0.376%	1.430%
470.2799042	489.8982322	480.089068	206	0.005345373	0.01964295	0.535%	1.964%
489.8982322	509.5165602	499.707396	235	0.006097877	0.025740827	0.610%	2.574%
509.5165602	529.1348882	519.325724	287	0.007447195	0.033188022	0.745%	3.319%
529.1348882	548.7532162	538.944052	371	0.009626862	0.042814884	0.963%	4.281%
548.7532162	568.3715442	558.56238	420	0.010898334	0.053713218	1.090%	5.371%
568.3715442	587.9898722	578.180708	526	0.013648866	0.067362084	1.365%	6.736%
587.9898722	607.6082003	597.799036	632	0.016399398	0.083761482	1.640%	8.376%
607.6082003	627.2265283	617.417364	715	0.018553116	0.102314599	1.855%	10.231%
627.2265283	646.8448563	637.035692	816	0.021173906	0.123488505	2.117%	12.349%
646.8448563	666.4631843	656.65402	843	0.021874513	0.145363018	2.187%	14.536%
666.4631843	686.0815123	676.272348	907	0.023535212	0.16889823	2.354%	16.890%
686.0815123	705.6998403	695.890676	954	0.024754787	0.193653018	2.475%	19.365%
705.6998403	725.3181683	715.509004	952	0.024702891	0.218355908	2.470%	21.836%
725.3181683	744.9364963	735.127332	1028	0.02667497	0.245030879	2.667%	24.503%
744.9364963	764.5548243	754.74566	1001	0.025974363	0.271005242	2.597%	27.101%
764.5548243	784.1731523	774.363988	971	0.025195911	0.296201152	2.520%	29.620%
784.1731523	803.7914804	793.982316	968	0.025118065	0.321319217	2.512%	32.132%
803.7914804	823.4098084	813.600644	910	0.023613057	0.344932275	2.361%	34.493%
823.4098084	843.0281364	833.218972	888	0.023042192	0.367974467	2.304%	36.797%
843.0281364	862.6464644	852.8373	859	0.022289688	0.390264155	2.229%	39.026%
862.6464644	882.2647924	872.455628	923	0.023950387	0.414214541	2.395%	41.421%
882.2647924	901.8831204	892.073956	819	0.021251752	0.435466293	2.125%	43.547%
901.8831204	921.5014484	911.692284	850	0.022056152	0.457522445	2.206%	45.752%

901.8831204	921.5014484	911.692284	850	17632	0.022056152	0.457522445	2.206%	45.752%
921.5014484	941.1197764	931.310612	821	18453	0.021303648	0.478826094	2.130%	47.883%
941.1197764	960.7381044	950.92894	767	19220	0.019902434	0.498728528	1.990%	49.873%
960.7381044	980.3564324	970.547268	752	19972	0.019513208	0.518241735	1.951%	51.824%
980.3564324	999.9747605	990.165596	764	20736	0.019824589	0.538066324	1.982%	53.807%
999.9747605	1019.593088	1009.78392	695	21431	0.018034148	0.556100472	1.803%	55.610%
1019.593088	1039.211416	1029.40225	660	22091	0.017125954	0.573226426	1.713%	57.323%
1039.211416	1058.829744	1049.02058	689	22780	0.017878458	0.591104883	1.788%	59.110%
1058.829744	1078.448073	1068.63891	712	23492	0.018475271	0.609580155	1.848%	60.958%
1078.448073	1098.066401	1088.25724	632	24124	0.016399398	0.625979553	1.640%	62.598%
1098.066401	1117.684729	1107.87556	644	24768	0.016710779	0.642690332	1.671%	64.269%
1117.684729	1137.303057	1127.49389	541	25309	0.014038092	0.656728424	1.404%	65.673%
1137.303057	1156.921385	1147.11222	574	25883	0.01489439	0.671622814	1.489%	67.162%
1156.921385	1176.539713	1166.73055	605	26488	0.015698791	0.687321605	1.570%	68.732%
1176.539713	1196.158041	1186.34888	547	27035	0.014193783	0.701515387	1.419%	70.152%
1196.158041	1215.776369	1205.9672	498	27533	0.01292231	0.714437698	1.292%	71.444%
1215.776369	1235.394697	1225.58553	542	28075	0.014064041	0.728501739	1.406%	72.850%
1235.394697	1255.013025	1245.20386	450	28525	0.011676787	0.740178525	1.168%	74.018%
1255.013025	1274.631353	1264.82219	457	28982	0.011858425	0.752036951	1.186%	75.204%
1274.631353	1294.249681	1284.44052	446	29428	0.011572993	0.763609943	1.157%	76.361%
1294.249681	1313.868009	1304.05884	421	29849	0.010924283	0.774534226	1.092%	77.453%
1313.868009	1333.486337	1323.67717	415	30264	0.010768592	0.785302818	1.077%	78.530%
1333.486337	1353.104665	1343.2955	403	30667	0.010457211	0.795760029	1.046%	79.576%
1353.104665	1372.722993	1362.91383	362	31029	0.009393326	0.805153355	0.939%	80.515%
1372.722993	1392.341321	1382.53216	345	31374	0.008952203	0.814105558	0.895%	81.411%
1392.341321	1411.959649	1402.15048	354	31728	0.009185739	0.823291297	0.919%	82.329%
1411.959649	1431.577977	1421.76881	321	32049	0.008329441	0.831620738	0.833%	83.162%
1431.577977	1451.196305	1441.38714	285	32334	0.007395298	0.839016036	0.740%	83.902%
1451.196305	1470.814633	1461.00547	310	32644	0.008044009	0.847060045	0.804%	84.706%
1470.814633	1490.432961	1480.6238	300	32944	0.007784524	0.854844569	0.778%	85.484%
1490.432961	1510.051289	1500.24212	291	33235	0.007550989	0.862395558	0.755%	86.240%
1510.051289	1529.669617	1519.86045	247	33482	0.006409258	0.868804816	0.641%	86.880%
1529.669617	1549.287945	1539.47878	227	33709	0.00589029	0.874695106	0.589%	87.470%
1549.287945	1568.906273	1559.09711	244	33953	0.006331413	0.881026519	0.633%	88.103%
1568.906273	1588.524601	1578.71544	218	34171	0.005656754	0.886683274	0.566%	88.668%
1588.524601	1608.142929	1598.33376	199	34370	0.005163734	0.891847008	0.516%	89.185%
1608.142929	1627.761257	1617.95209	187	34557	0.004852354	0.896699362	0.485%	89.670%
1627.761257	1647.379585	1637.57042	194	34751	0.005033992	0.901733354	0.503%	90.173%
1647.379585	1666.997913	1657.18875	171	34922	0.004437179	0.906170533	0.444%	90.617%
1666.997913	1686.616241	1676.80708	159	35081	0.004125798	0.910296331	0.413%	91.030%
1686.616241	1706.234569	1696.4254	147	35228	0.003814417	0.914110748	0.381%	91.411%
1706.234569	1725.852897	1716.04373	132	35360	0.003425191	0.917535939	0.343%	91.754%
1725.852897	1745.471225	1735.66206	156	35516	0.004047953	0.921583891	0.405%	92.158%
1745.471225	1765.089553	1755.28039	129	35645	0.003347345	0.924931237	0.335%	92.493%
1765.089553	1784.707881	1774.89872	110	35755	0.002854326	0.927785562	0.285%	92.779%
1784.707881	1804.326209	1794.51704	160	35915	0.004151746	0.931937309	0.415%	93.194%
1804.326209	1823.944537	1814.13537	124	36039	0.003217603	0.935154912	0.322%	93.515%
1823.944537	1843.562865	1833.7537	110	36149	0.002854326	0.938009238	0.285%	93.801%
1843.562865	1863.181193	1853.37203	104	36253	0.002698635	0.940707873	0.270%	94.071%
1863.181193	1882.799521	1872.99036	95	36348	0.002465099	0.943172972	0.247%	94.317%
1882.799521	1902.417849	1892.60868	111	36459	0.002880274	0.946053246	0.288%	94.605%

1902.417849	1922.036177	1912.22701	110	36569	0.002854326	0.948907572	0.285%	94.891%
1922.036177	1941.654505	1931.84534	92	36661	0.002387254	0.951294826	0.239%	95.129%
1941.654505	1961.272833	1951.46367	70	36731	0.001816389	0.953111215	0.182%	95.311%
1961.272833	1980.891161	1971.082	101	36832	0.00262079	0.955732005	0.262%	95.573%
1980.891161	2000.509489	1990.70032	82	36914	0.00212777	0.957859775	0.213%	95.786%
2000.509489	2020.127817	2010.31865	99	37013	0.002568893	0.960428668	0.257%	96.043%
2020.127817	2039.746145	2029.93698	64	37077	0.001660699	0.962089366	0.166%	96.209%
2039.746145	2059.364473	2049.55531	93	37170	0.002413203	0.964502569	0.241%	96.450%
2059.364473	2078.982801	2069.17364	63	37233	0.00163475	0.966137319	0.163%	96.614%
2078.982801	2098.601129	2088.79197	58	37291	0.001505008	0.967642327	0.151%	96.764%
2098.601129	2118.219457	2108.41029	58	37349	0.001505008	0.969147335	0.151%	96.915%
2118.219457	2137.837785	2128.02862	57	37406	0.00147906	0.970626395	0.148%	97.063%
2137.837785	2157.456113	2147.64695	50	37456	0.001297421	0.971923815	0.130%	97.192%
2157.456113	2177.074441	2167.26528	56	37512	0.001453111	0.973376927	0.145%	97.338%
2177.074441	2196.692769	2186.88361	62	37574	0.001608802	0.974985728	0.161%	97.499%
2196.692769	2216.311097	2206.50193	51	37625	0.001323369	0.976309098	0.132%	97.631%
2216.311097	2235.929425	2226.12026	40	37665	0.001037937	0.977347034	0.104%	97.735%
2235.929425	2255.547753	2245.73859	52	37717	0.001349318	0.978696352	0.135%	97.870%
2255.547753	2275.166081	2265.35692	49	37766	0.001271472	0.979967824	0.127%	97.997%
2275.166081	2294.784409	2284.97525	35	37801	0.000908195	0.980876018	0.091%	98.088%
2294.784409	2314.402737	2304.59357	26	37827	0.000674659	0.981550677	0.067%	98.155%
2314.402737	2334.021065	2324.2119	30	37857	0.000778452	0.98232913	0.078%	98.233%
2334.021065	2353.639393	2343.83023	37	37894	0.000960091	0.983289221	0.096%	98.329%
2353.639393	2373.257721	2363.44856	29	37923	0.000752504	0.984041725	0.075%	98.404%
2373.257721	2392.876049	2383.06689	35	37958	0.000908195	0.98494992	0.091%	98.495%
2392.876049	2412.494377	2402.68521	18	37976	0.000467071	0.985416991	0.047%	98.542%
2412.494377	2432.112705	2422.30354	26	38002	0.000674659	0.98609165	0.067%	98.609%
2432.112705	2451.731033	2441.92187	26	38028	0.000674659	0.986766309	0.067%	98.677%
2451.731033	2471.349361	2461.5402	22	38050	0.000570865	0.987337174	0.057%	98.734%
2471.349361	2490.967689	2481.15853	19	38069	0.00049302	0.987830194	0.049%	98.783%
2490.967689	2510.586017	2500.77685	30	38099	0.000778452	0.988608646	0.078%	98.861%
2510.586017	2530.204345	2520.39518	30	38129	0.000778452	0.989387098	0.078%	98.939%
2530.204345	2549.822673	2540.01351	18	38147	0.000467071	0.98985417	0.047%	98.985%
2549.822673	2569.441001	2559.63184	21	38168	0.000544917	0.990399087	0.054%	99.040%
2569.441001	2589.059329	2579.25017	23	38191	0.000596814	0.9909959	0.060%	99.100%
2589.059329	2608.677657	2598.86849	22	38213	0.000570865	0.991566765	0.057%	99.157%
2608.677657	2628.295985	2618.48682	19	38232	0.00049302	0.992059785	0.049%	99.206%
2628.295985	2647.914313	2638.10515	17	38249	0.000441123	0.992500908	0.044%	99.250%
2647.914313	2667.532641	2657.72348	17	38266	0.000441123	0.992942031	0.044%	99.294%
2667.532641	2687.150969	2677.34181	11	38277	0.000285433	0.993227464	0.029%	99.323%
2687.150969	2706.769297	2696.96013	17	38294	0.000441123	0.993668587	0.044%	99.367%
2706.769297	2726.387625	2716.57846	11	38305	0.000285433	0.993954019	0.029%	99.395%
2726.387625	2746.005953	2736.19679	9	38314	0.000233536	0.994187555	0.023%	99.419%
2746.005953	2765.624281	2755.81512	8	38322	0.000207587	0.994395142	0.021%	99.440%
2765.624281	2785.242609	2775.43345	18	38340	0.000467071	0.994862214	0.047%	99.486%
2785.242609	2804.860937	2795.05177	11	38351	0.000285433	0.995147646	0.029%	99.515%
2804.860937	2824.479265	2814.6701	7	38358	0.000181639	0.995329285	0.018%	99.533%
2824.479265	2844.097593	2834.28843	8	38366	0.000207587	0.995536873	0.021%	99.554%
2844.097593	2863.715921	2853.90676	11	38377	0.000285433	0.995822305	0.029%	99.582%
2863.715921	2883.334249	2873.52509	9	38386	0.000233536	0.996055841	0.023%	99.606%

2883.334249	2902.952577	2893.14341	10	38396	0.000259484	0.996315325	0.026%	99.632%
2902.952577	2922.570905	2912.76174	7	38403	0.000181639	0.996496964	0.018%	99.650%
2922.570905	2942.189233	2932.38007	7	38410	0.000181639	0.996678603	0.018%	99.668%
2942.189233	2961.807561	2951.9984	4	38414	0.000103794	0.996782397	0.010%	99.678%
2961.807561	2981.425889	2971.61673	3	38417	7.78452E-05	0.996860242	0.008%	99.686%
2981.425889	3001.044218	2991.23505	10	38427	0.000259484	0.997119726	0.026%	99.712%
3001.044218	3020.662546	3010.85338	5	38432	0.000129742	0.997249468	0.013%	99.725%
3020.662546	3040.280874	3030.47171	9	38441	0.000233536	0.997483004	0.023%	99.748%
3040.280874	3059.899202	3050.09004	4	38445	0.000103794	0.997586797	0.010%	99.759%
3059.899202	3079.51753	3069.70837	4	38449	0.000103794	0.997690591	0.010%	99.769%
3079.51753	3099.135858	3089.32669	6	38455	0.00015569	0.997846282	0.016%	99.785%
3099.135858	3118.754186	3108.94502	5	38460	0.000129742	0.997976024	0.013%	99.798%
3118.754186	3138.372514	3128.56335	5	38465	0.000129742	0.998105766	0.013%	99.811%
3138.372514	3157.990842	3148.18168	5	38470	0.000129742	0.998235508	0.013%	99.824%
3157.990842	3177.60917	3167.80001	6	38476	0.00015569	0.998391198	0.016%	99.839%
3177.60917	3197.227498	3187.41833	1	38477	2.59484E-05	0.998417147	0.003%	99.842%
3197.227498	3216.845826	3207.03666	3	38480	7.78452E-05	0.998494992	0.008%	99.849%
3216.845826	3236.464154	3226.65499	2	38482	5.18968E-05	0.998546889	0.005%	99.855%
3236.464154	3256.082482	3246.27332	4	38486	0.000103794	0.998650682	0.010%	99.865%
3256.082482	3275.70081	3265.89165	1	38487	2.59484E-05	0.998676631	0.003%	99.868%
3275.70081	3295.319138	3285.50997	4	38491	0.000103794	0.998780425	0.010%	99.878%
3295.319138	3314.937466	3305.1283	2	38493	5.18968E-05	0.998832321	0.005%	99.883%
3314.937466	3334.555794	3324.74663	0	38493	0	0.998832321	0.000%	99.883%
3334.555794	3354.174122	3344.36496	2	38495	5.18968E-05	0.998884218	0.005%	99.888%
3354.174122	3373.79245	3363.98329	8	38503	0.000207587	0.999091805	0.021%	99.909%
3373.79245	3393.410778	3383.60161	1	38504	2.59484E-05	0.999117754	0.003%	99.912%
3393.410778	3413.029106	3403.21994	2	38506	5.18968E-05	0.999169651	0.005%	99.917%
3413.029106	3432.647434	3422.83827	0	38506	0	0.999169651	0.000%	99.917%
3432.647434	3452.265762	3442.4566	4	38510	0.000103794	0.999273444	0.010%	99.927%
3452.265762	3471.88409	3462.07493	3	38513	7.78452E-05	0.99935129	0.008%	99.935%
3471.88409	3491.502418	3481.69325	0	38513	0	0.99935129	0.000%	99.935%
3491.502418	3511.120746	3501.31158	0	38513	0	0.99935129	0.000%	99.935%
3511.120746	3530.739074	3520.92991	1	38514	2.59484E-05	0.999377238	0.003%	99.938%
3530.739074	3550.357402	3540.54824	3	38517	7.78452E-05	0.999455083	0.008%	99.946%
3550.357402	3569.97573	3560.16657	0	38517	0	0.999455083	0.000%	99.946%
3569.97573	3589.594058	3579.78489	2	38519	5.18968E-05	0.99950698	0.005%	99.951%
3589.594058	3609.212386	3599.40322	1	38520	2.59484E-05	0.999532929	0.003%	99.953%
3609.212386	3628.830714	3619.02155	1	38521	2.59484E-05	0.999558877	0.003%	99.956%
3628.830714	3648.449042	3638.63988	4	38525	0.000103794	0.999662671	0.010%	99.966%
3648.449042	3668.06737	3658.25821	1	38526	2.59484E-05	0.999688619	0.003%	99.969%
3668.06737	3687.685698	3677.87653	0	38526	0	0.999688619	0.000%	99.969%
3687.685698	3707.304026	3697.49486	2	38528	5.18968E-05	0.999740516	0.005%	99.974%
3707.304026	3726.922354	3717.11319	0	38528	0	0.999740516	0.000%	99.974%
3726.922354	3746.540682	3736.73152	0	38528	0	0.999740516	0.000%	99.974%
3746.540682	3766.15901	3756.34985	0	38528	0	0.999740516	0.000%	99.974%
3766.15901	3785.777338	3775.96817	1	38529	2.59484E-05	0.999766464	0.003%	99.977%
3785.777338	3805.395666	3795.5865	0	38529	0	0.999766464	0.000%	99.977%
3805.395666	3825.013994	3815.20483	0	38529	0	0.999766464	0.000%	99.977%
3825.013994	3844.632322	3834.82316	0	38529	0	0.999766464	0.000%	99.977%
3844.632322	3864.25065	3854.44149	0	38529	0	0.999766464	0.000%	99.977%
3864.25065	3883.868978	3874.05981	2	38531	5.18968E-05	0.999818361	0.005%	99.982%
3883.868978	3903.487306	3893.67814	0	38531	0	0.999818361	0.000%	99.982%
3903.487306	3923.105634	3913.29647	0	38531	0	0.999818361	0.000%	99.982%
3923.105634	3942.723962	3932.9148	0	38531	0	0.999818361	0.000%	99.982%
3942.723962	3962.34229	3952.53313	7	38538	0.000181639	1	0.018%	100.000%

Bibliografía.

- Rincón, L. (2019, 5 octubre). *Una introducción a la estadística inferencial*. UNAM. Recuperado 6 de noviembre de 2021, de <https://lya.fciencias.unam.mx/lars/Publicaciones/ei2019.pdf>
- K. (2012, 28 julio). *Estadística inferencial*. Slideshare. Recuperado 6 de noviembre de 2021, de <https://es.slideshare.net/katemora/proyecto-estadistica-inferencial-13786749>
- Moreno, J. (2020, 14 septiembre). *Golden Oaks, Florida. ¿Tu casa en Walt Disney World? El Souvenir*. Recuperado 5 de noviembre de 2021, de <https://elsouvenir.com/golden-oaks-florida-tu-casa-en-walt-disney-world/>
- Wikipedia contributors. (2021, 12 agosto). *Golden Oak at Walt Disney World Resort*. Wikipedia. Recuperado 5 de noviembre de 2021, de https://en.wikipedia.org/wiki/Golden_Oak_at_Walt_Disney_World_Resort