

# Proyecto (segundo parcial): análisis de una empresa de entretenimiento

CARLOS ENRIQUE LEZAMA JACINTO (181121) and PAULINA GUDALUPE MONTIEL GALLARDO (183629), Instituto Tecnológico Autónomo de México, México

Como evaluación para el segundo parcial de la materia de Inferencia Estadística se realizará un análisis de una empresa de entretenimiento en Estados Unidos con el objetivo de brindar, con base en resultados estadísticos, recomendaciones acerca del manejo de esta organización. Para ello, se realizaron diversos cálculos en el lenguaje R, con el fin de optimizar los mismos.

## 1. PLANTEAMIENTO DEL PROBLEMA

Una cadena norteamericana dedicada al entretenimiento posee varias salas de cine en la costa Este de los Estados Unidos. La empresa desea llevar a cabo un análisis con la información recopilada en los últimos dos años y medio.

El detalle de los datos con los que cuenta dicha cadena es a nivel semanal y por región. Como las películas suelen estrenarse en viernes para maximizar la cantidad de gente que acude a verlas, la logística de la empresa maneja semanas que empiezan en viernes y terminan en jueves. Por otro lado, las regiones en las que está dividida la operación son cuatro: Nueva Inglaterra, Atlántico Central, Sureste y Florida (respectivamente 1, 2, 3 y 4 en las bases de datos).

La información proveniente de la dulcería de los cines cuenta con dos variables: la venta total en dólares por tipo de producto, así como el número total de productos vendidos. Los tipos de producto se dividen en lo que la empresa denomina familias: palomitas, bebidas embotelladas, bebidas de cafetería, combos (palomitas con bebidas) y otros (como postres, nuggets, etc).

Adicionalmente, la información proveniente de taquillas cuenta con los datos del total de boletos vendidos durante cada semana, así como el número de transacciones realizadas. Considere que una transacción puede contener 1, 2 o hasta  $k$  boletos.

Por último, se cuenta con el dato, para cada semana, de si hubo algún estreno de largometrajes considerados blockbusters. La compañía considera que una película es un blockbuster si ésta recaudó, durante todo el tiempo que estuvo en cartelera, más de 15 millones de dólares.

### 1.1. Problemas a resolver

- ¿Han disminuido las ventas de palomitas, bebidas y combos en los últimos años?
- ¿El comportamiento de las ventas en dulcería es similar en las 4 zonas?
- ¿La venta individual de palomitas y bebidas no es relevante y, por lo tanto, es mejor enfocarse únicamente en la venta de combos?
- ¿Es posible determinar que, mientras mayor sea el número de visitantes por semana, mayores son las ventas en dulcería?

---

Dirección de los autores: Carlos Enrique Lezama Jacinto (181121), clezamaj@itam.mx; Paulina Guadalupe Montiel Gallardo (183629), pmontie3@itam.mx, Instituto Tecnológico Autónomo de México, Río Hondo, 1, Álvaro Obregón, Ciudad de México, México, 01080.

---

2. CALCULA COMO NUEVAS VARIABLES EL PRECIO PROMEDIO POR TIPO DE PRODUCTO Y EL NÚMERO PROMEDIO DE VISITANTES A LA SEMANA.

Familia	Precio promedio
Bebidas embotelladas	\$ 28.10
Bebidas de la cafetería	\$ 57.80
Combos	\$ 132.77
Otros	\$ 54.03
Palomitas	\$ 59.77

Cuadro 1. Precios promedio por tipo de producto

Número promedio de visitantes a la semana: 3 820 207.16

3. CONSIDERANDO ESTAS NUEVAS VARIABLES, REALIZA UN ANÁLISIS EXPLORATORIO TANTO CON LA BASE DE ALIMENTOS COMO CON LA VENTA DE TAQUILLA:
- 3.1. Realiza histogramas y diagramas de caja y brazos para los datos por Zona y a total país. ¿Qué diferencias observas? Calcula estadísticos básicos como media, mediana, cuartiles, desviación y coeficiente de variación. ¿Qué puedes concluir a partir de toda esta información?

Variable	$\bar{X}$	$q_1$	$q_2$	$q_3$	$s$	$\hat{c}_v$
Ventas en dulcería	\$ 4 480 556.83	\$ 269 366.25	\$ 2 416 917.44	\$ 7 789 327.30	\$ 5 004 988.21	1.1170
Precio	\$ 66.80	\$ 49.65	\$ 56.46	\$ 64.74	\$ 36.68	0.5492
Venta de boletos	955 051.79	733 113	893 457	1 112 295.75	315 348.73	0.3302

Cuadro 2. Estadísticos básicos del total de la muestra

Variable	$\bar{X}$	$q_1$	$q_2$	$q_3$	$s$	$\hat{c}_v$
Ventas en dulcería	\$ 4 876 878.74	\$ 286 431.00	\$ 2 909 610.08	\$ 8 241 570.90	\$ 5 433 298.61	1.1141
Precio	\$ 65.65	\$ 47.87	\$ 55.90	\$ 63.94	\$ 35.92	0.5471
Venta de boletos	1 066 678.8	836 164.5	1 012 348	1 204 269.5	314 083.83	0.2944

Cuadro 3. Estadísticos básicos de la Zona 1 (Nueva Inglaterra)

Variable	$\bar{X}$	$q_1$	$q_2$	$q_3$	$s$	$\hat{c}_v$
Ventas en dulcería	\$ 3 711 948.27	\$ 202 881.50	\$ 1 808 397.05	\$ 6 430 021.21	\$ 4 222 280.43	1.1374
Precio	\$ 67.36	\$ 51.47	\$ 56.61	\$ 64.16	\$ 37.11	0.5509
Venta de boletos	750 246.23	585 362.5	713 332	853 630.5	218 673.60	0.2914

Cuadro 4. Estadísticos básicos de la Zona 2 (Atlántico Central)

Variable	$\bar{X}$	$q_1$	$q_2$	$q_3$	$s$	$\widehat{c}_v$
Ventas en dulcería	\$ 4 482 400.31	\$ 2 94 221.50	\$ 2 306 396.92	\$ 8 063 824.14	\$ 4 855 185.60	1.0831
Precio	\$ 66.80	\$ 49.65	\$ 56.46	\$ 64.74	\$ 36.68	0.5492
Venta de boletos	939 729.37	723 108	900 006	1 079 593.5	301 902.04	0.3212

Cuadro 5. Estadísticos básicos de la Zona 3 (Sureste)

Variable	$\bar{X}$	$q_1$	$q_2$	$q_3$	$s$	$\widehat{c}_v$
Ventas en dulcería	\$ 4 850 999.97	\$ 258 094.00	\$ 2 573 214.90	\$ 8 573 123.06	\$ 5 338 870.49	1.1005
Precio	\$ 64.95	\$ 48.97	\$ 55.04	\$ 63.08	\$ 35.56	0.5475
Venta de boletos	1 063 552.74	824 991.5	1 029 232	1 205 042	3 09 006.89	0.2905

Cuadro 6. Estadísticos básicos de la Zona 4 (Florida)

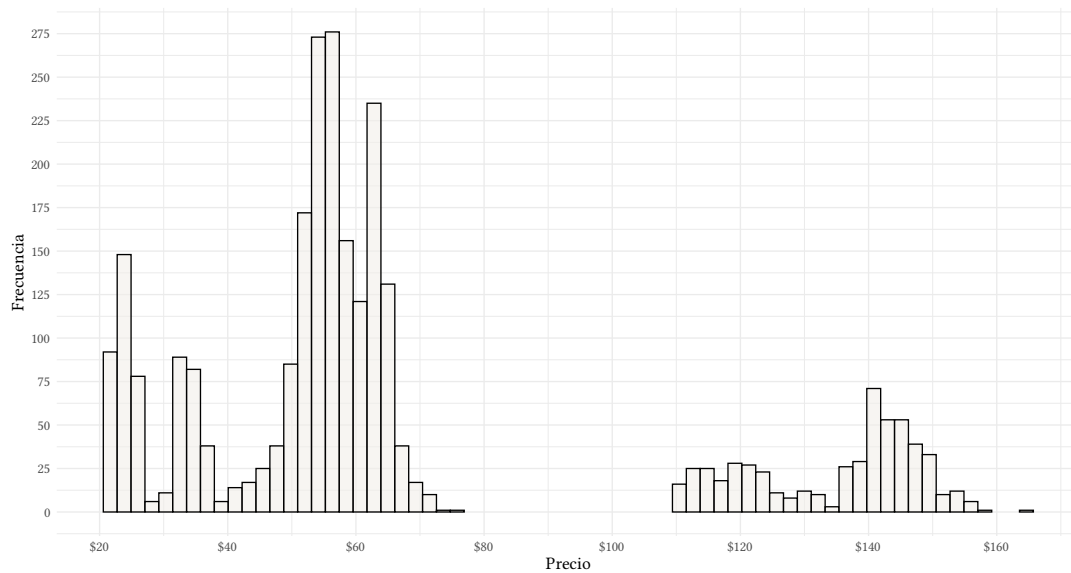


Fig. 1. Precios de dulcería

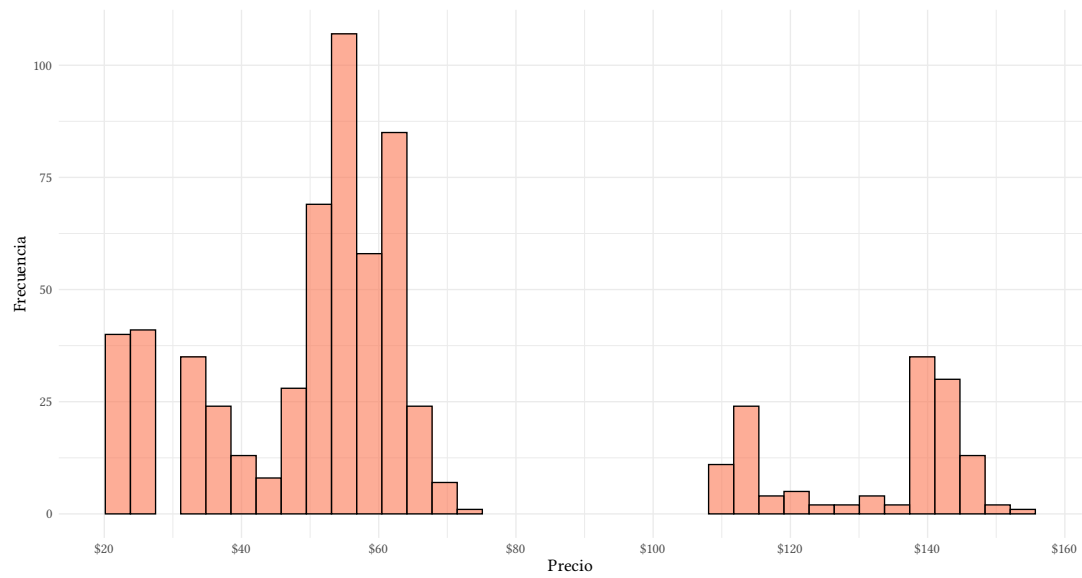


Fig. 2. Precios de dulcería en Nueva Inglaterra

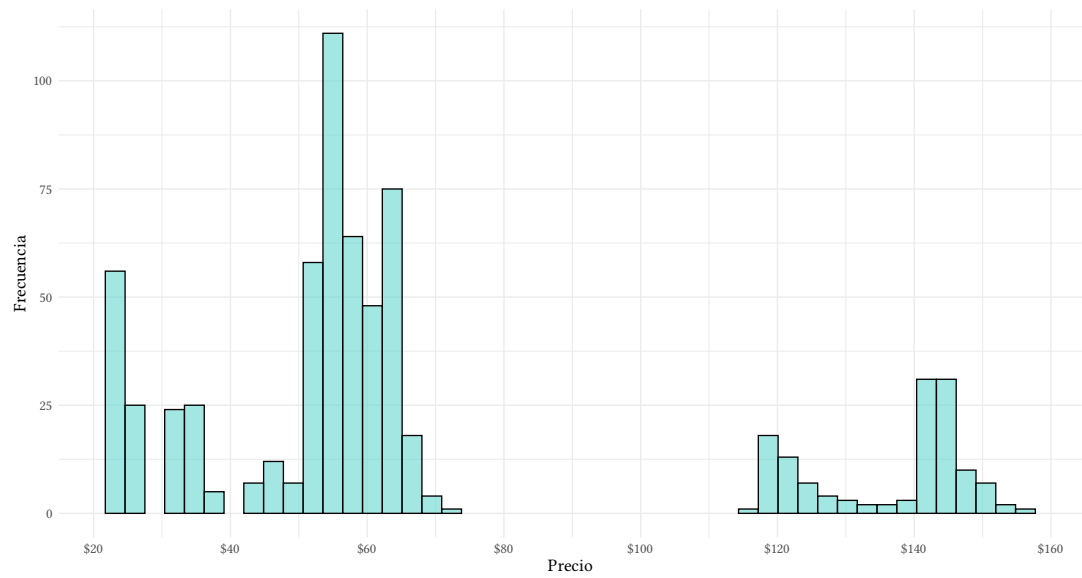


Fig. 3. Precios de dulcería en el Atlántico Central

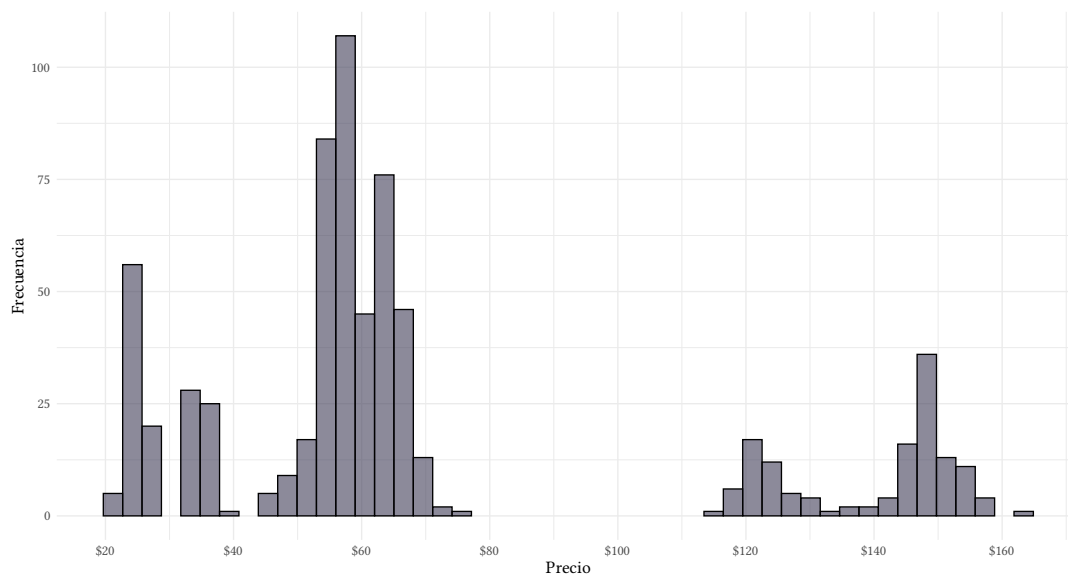


Fig. 4. Precios de dulcería en el Sureste

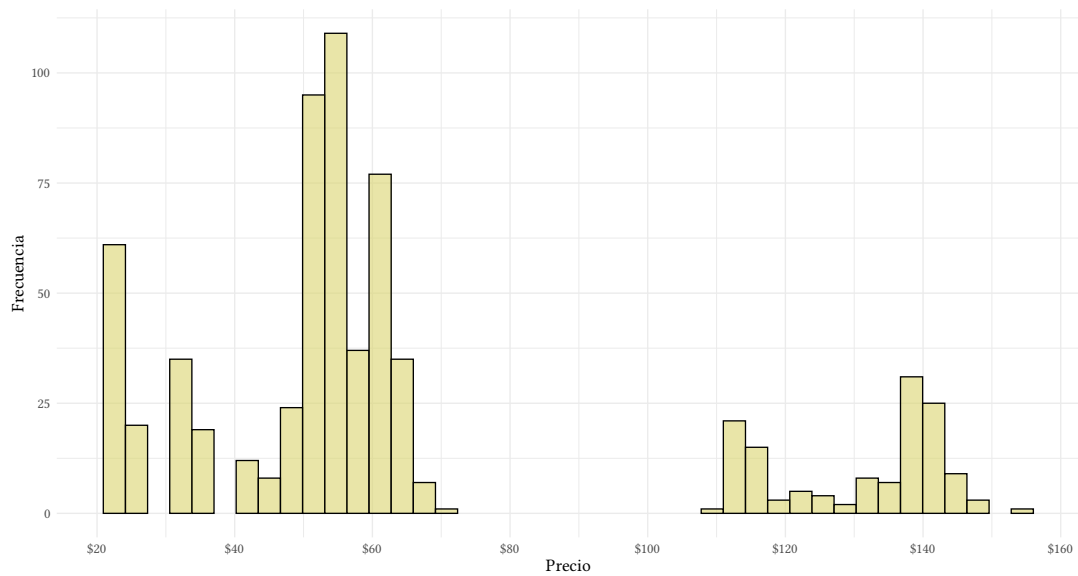


Fig. 5. Precios de dulcería en Florida

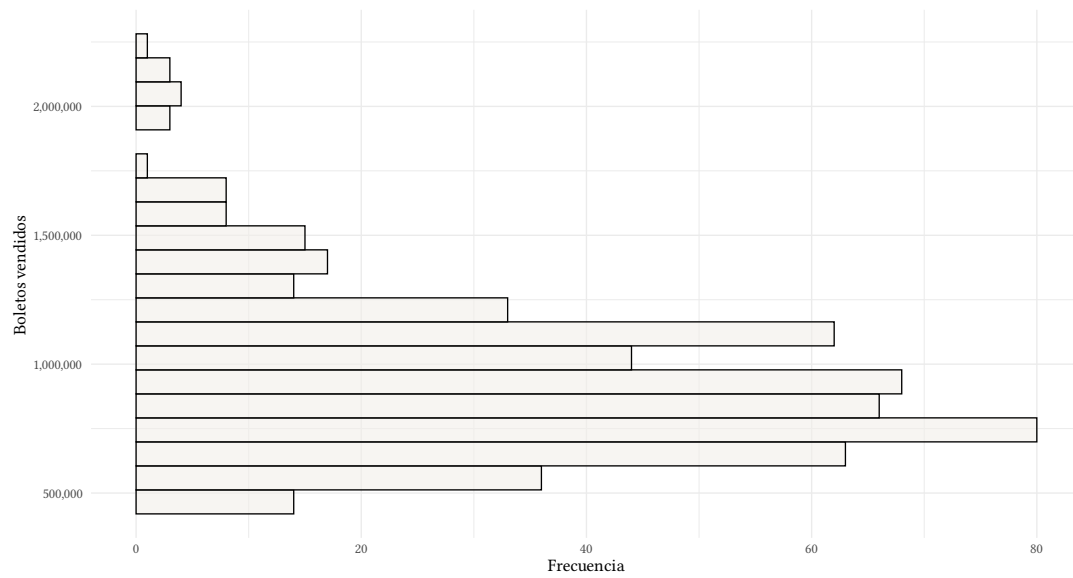


Fig. 6. Venta de boletos

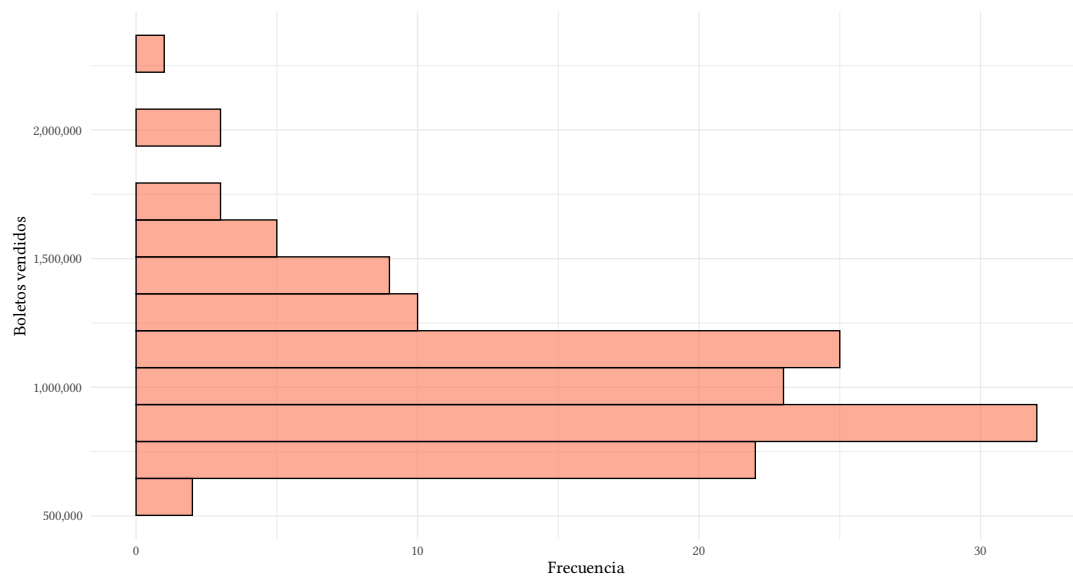


Fig. 7. Venta de boletos en Nueva Inglaterra

Fecha: Noviembre 2020.

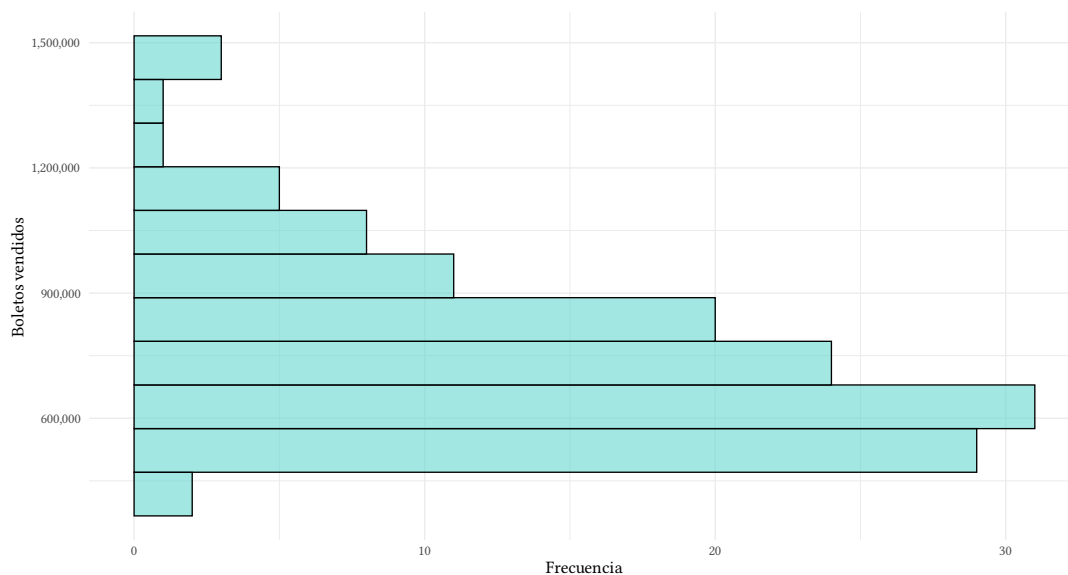


Fig. 8. Venta de boletos en el Atlántico Central

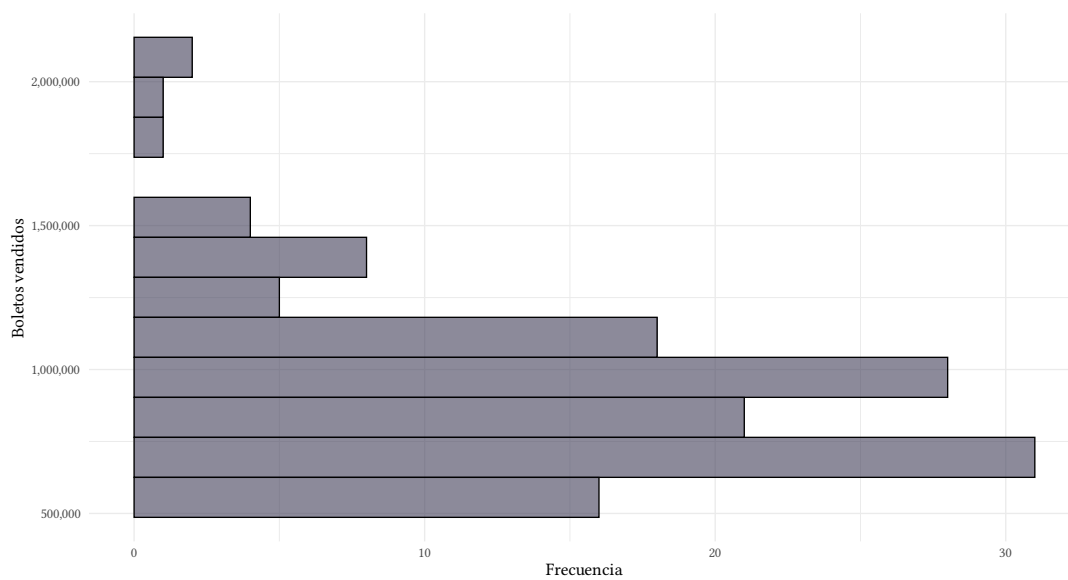


Fig. 9. Venta de boletos en el Sureste

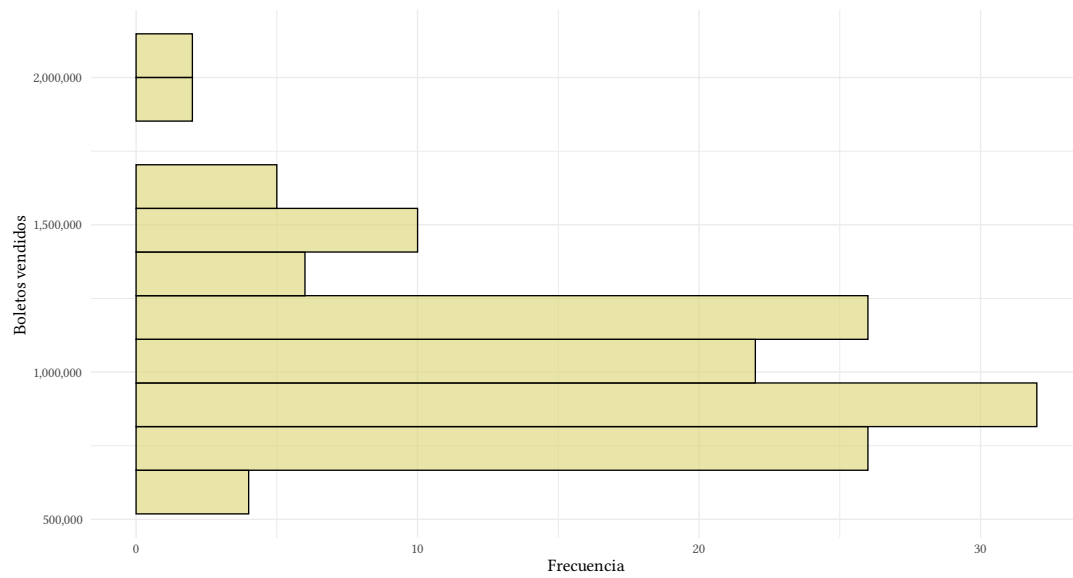


Fig. 10. Venta de boletos en Florida

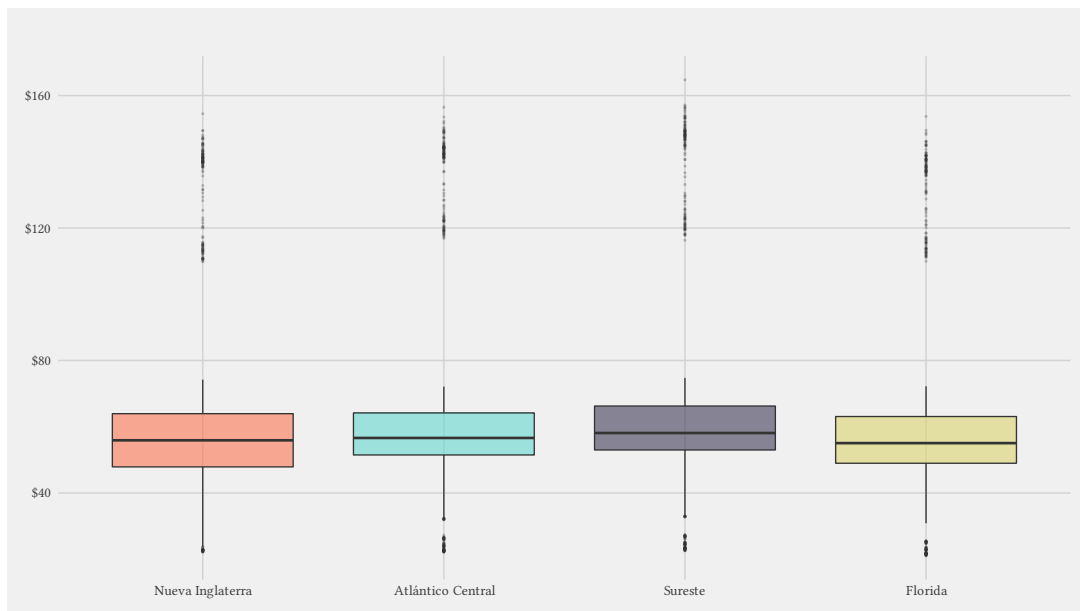


Fig. 11. Precios de dulcería por Zona



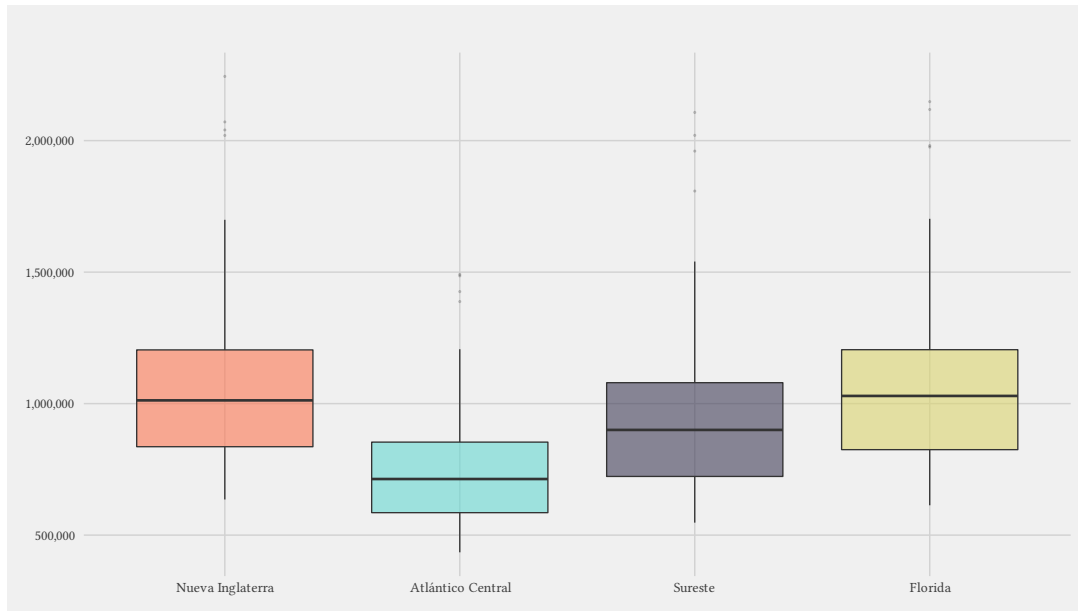


Fig. 12. Venta de boletos por Zona

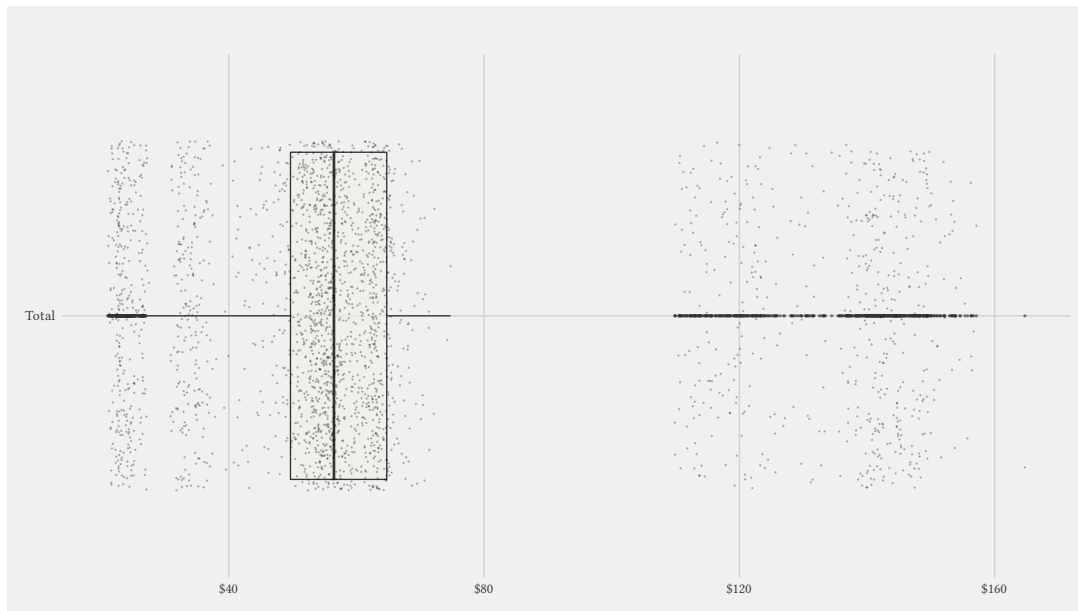


Fig. 13. Precios de dulcería

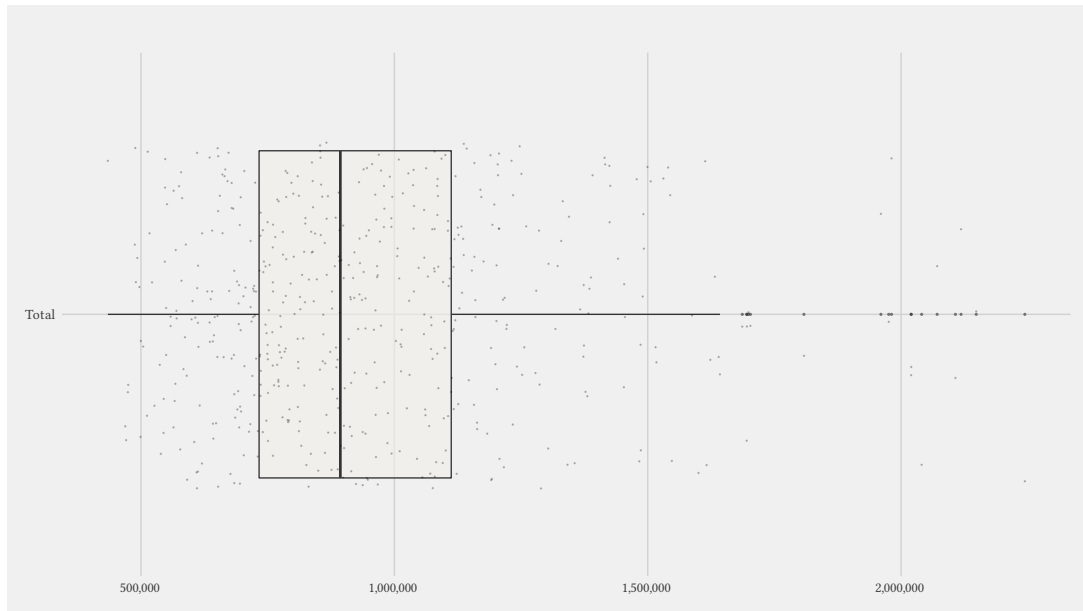


Fig. 14. Venta de boletos

### 3.2. Grafica las ventas a total y por Zona a lo largo del tiempo. ¿Qué patrones observas?

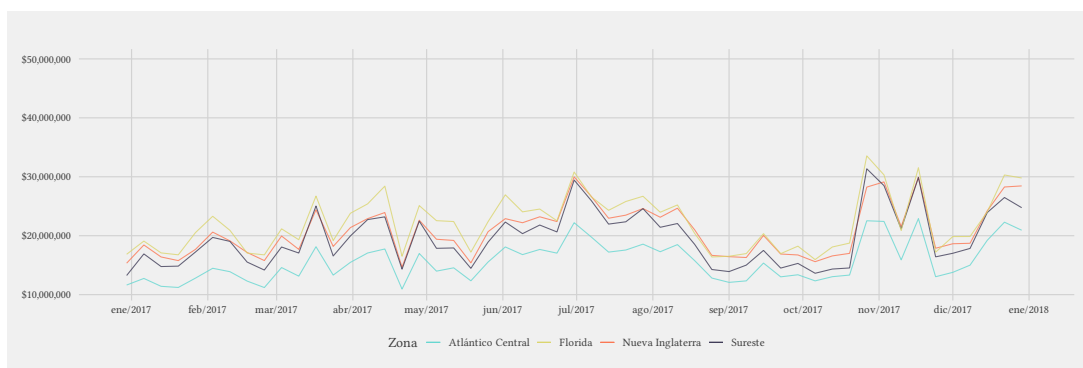


Fig. 15. Ventas en dulcería de 2017

Fecha: Noviembre 2020.

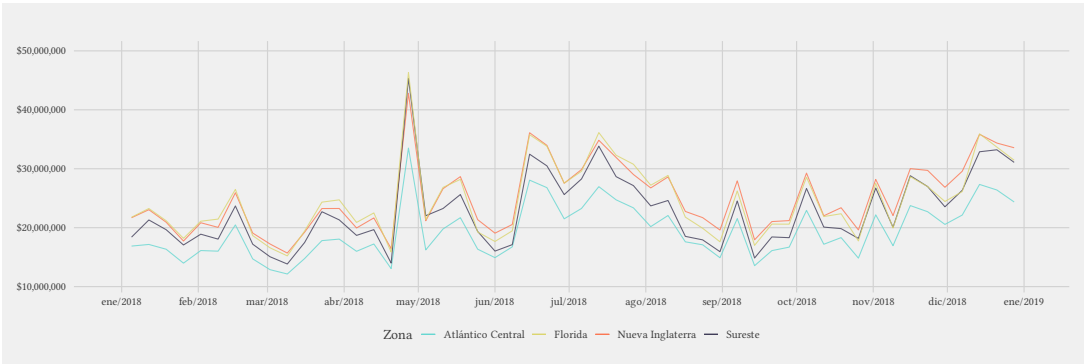


Fig. 16. Ventas en dulcería de 2018

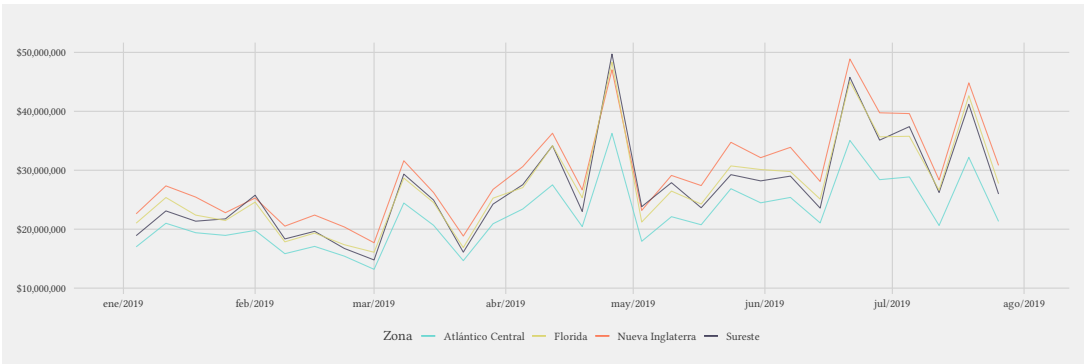


Fig. 17. Ventas en dulcería de 2019

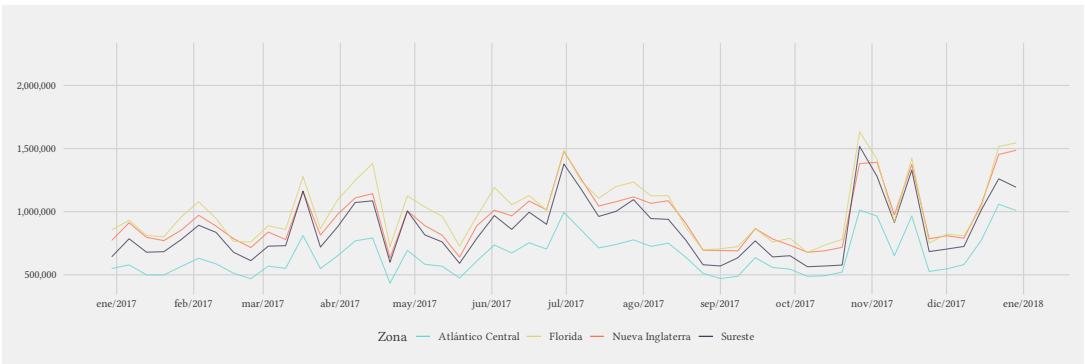


Fig. 18. Venta de boletos en 2017

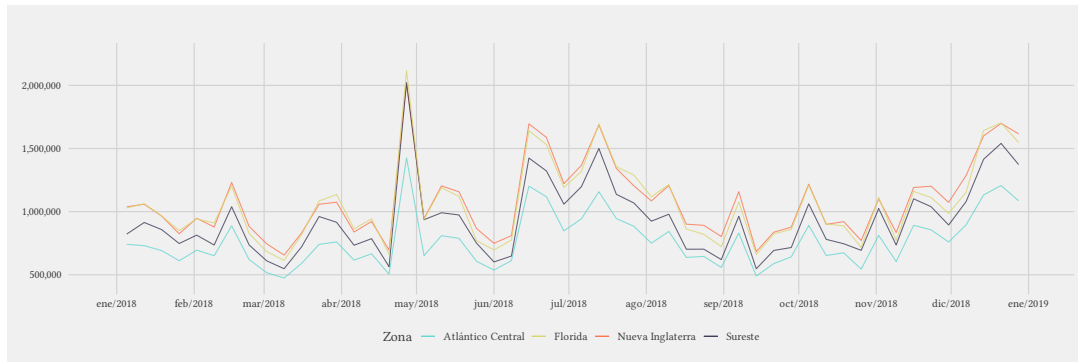


Fig. 19. Venta de boletos en 2018

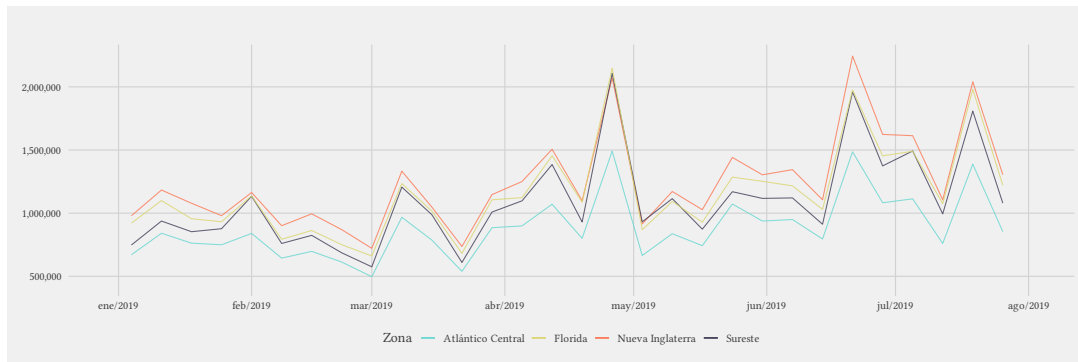


Fig. 20. Venta de boletos en 2019

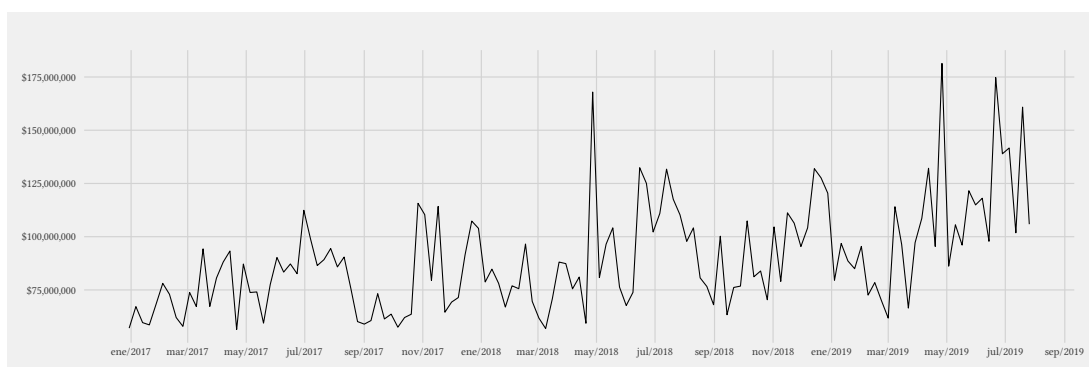


Fig. 21. Ventas en dulcería

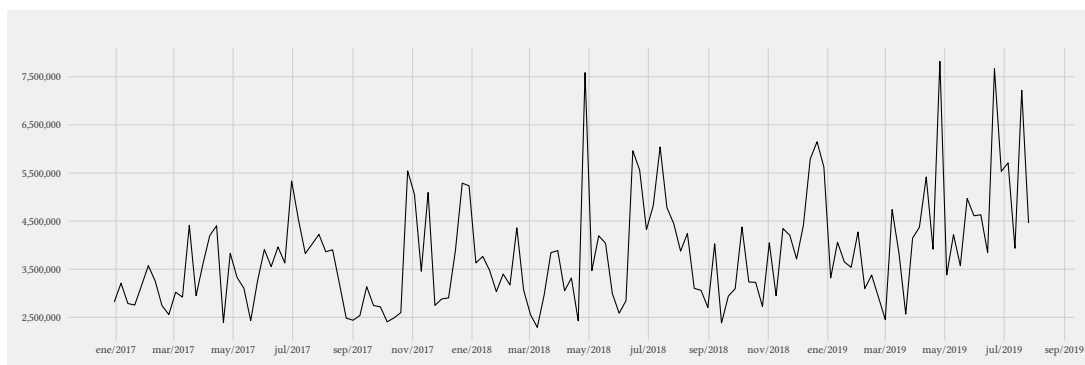


Fig. 22. Venta de boletos

### 3.3. Junta las bases considerando como llave la *semana*.

Obtén una medida de asociación lineal entre:

- Ventas en cada tipo de producto y el número promedio de asistentes.
- Precio de los productos y el número promedio de asistentes.

¿Qué puedes concluir al respecto?

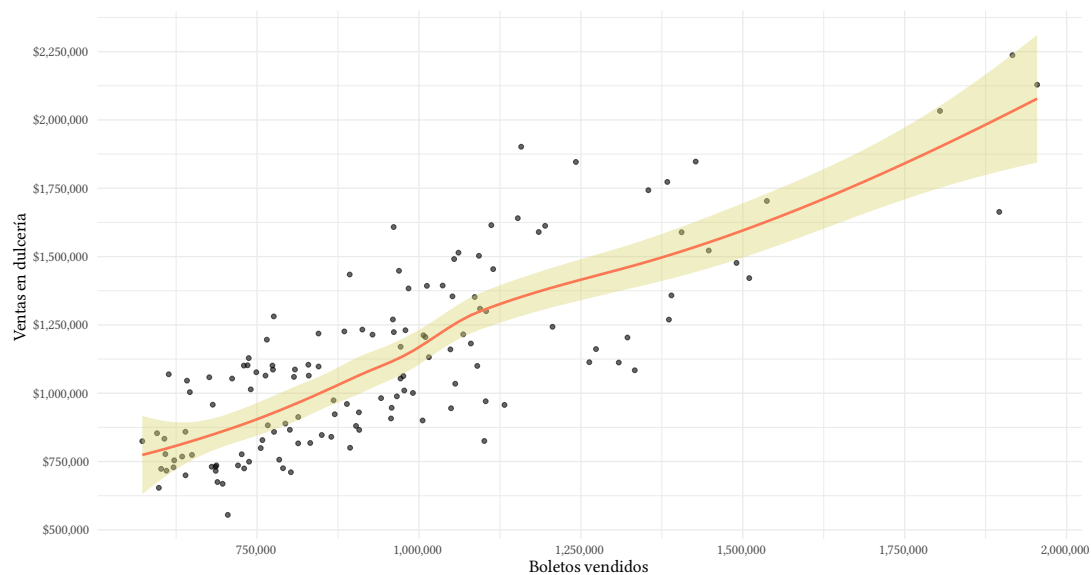


Fig. 23. Asistentes semanales promedio v. venta de bebidas embotelladas

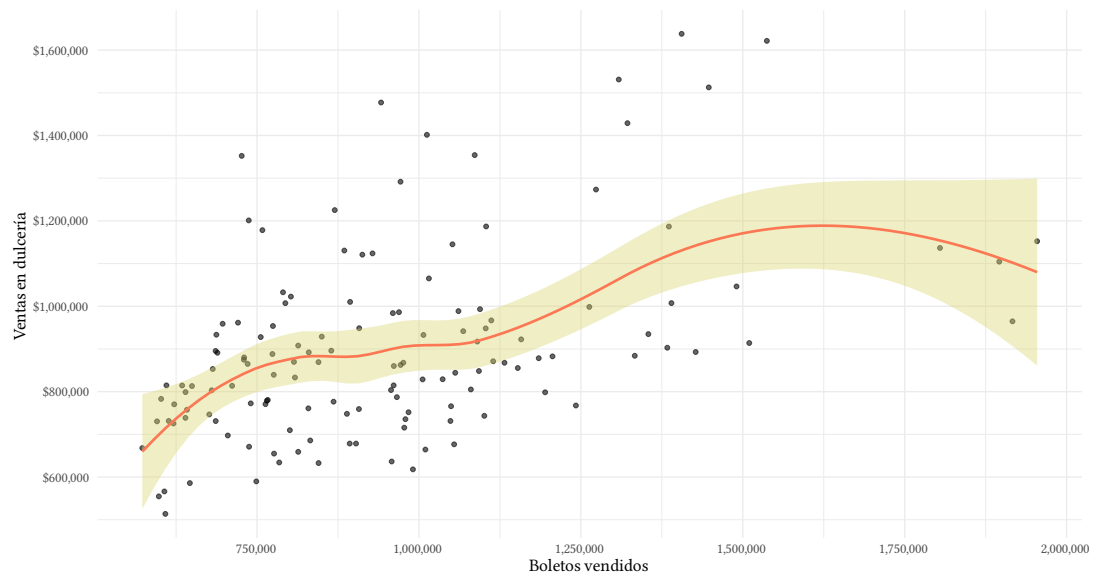


Fig. 24. Asistentes semanales promedio v. venta de bebidas de cafetería

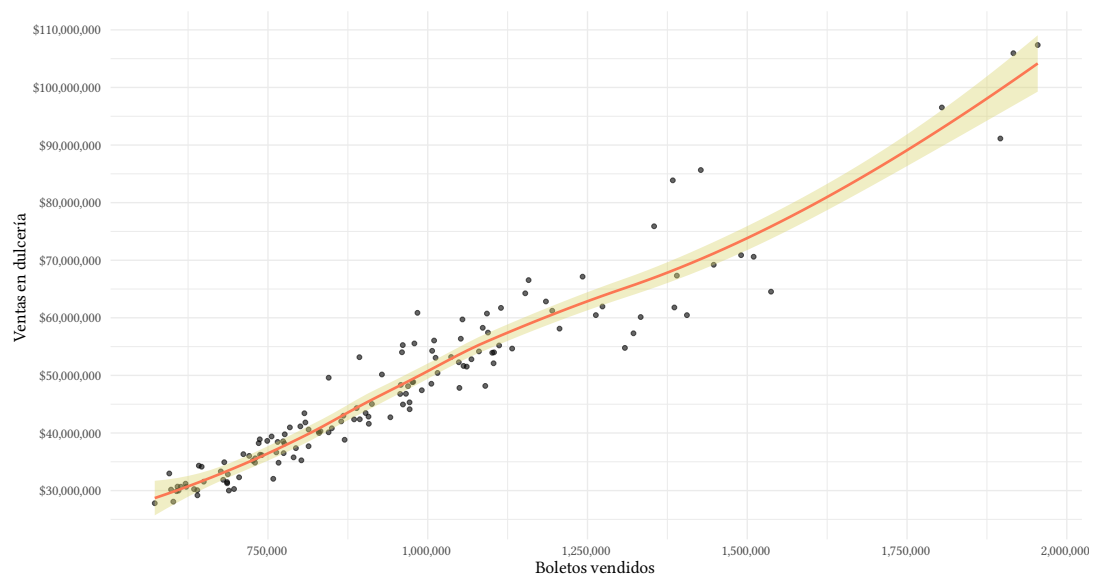


Fig. 25. Asistentes semanales promedio v. venta de combos

Fecha: Noviembre 2020.

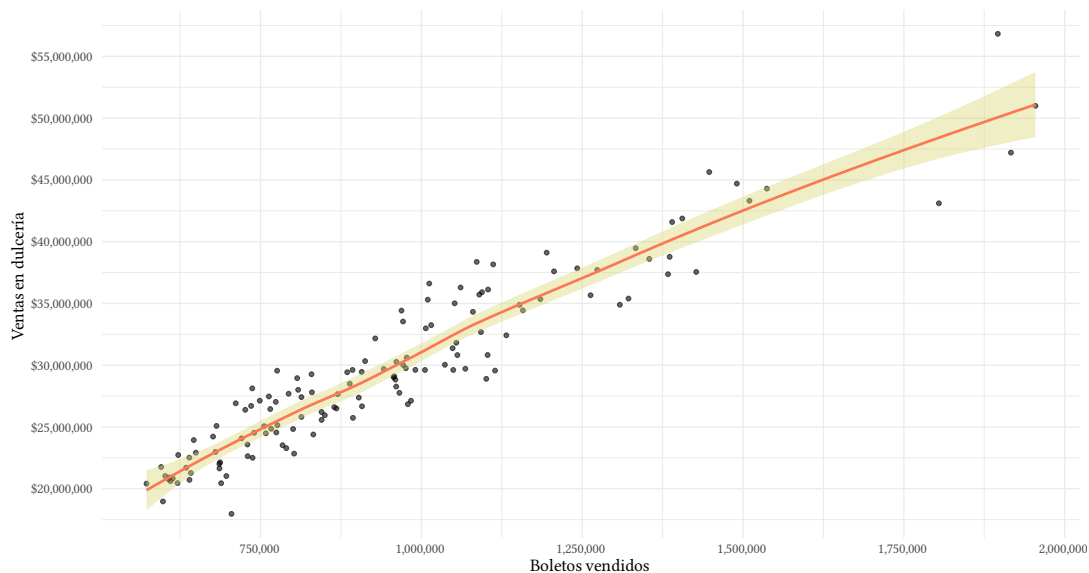


Fig. 26. Asistentes semanales promedio v. venta de postres, nachos, nuggets, etc.

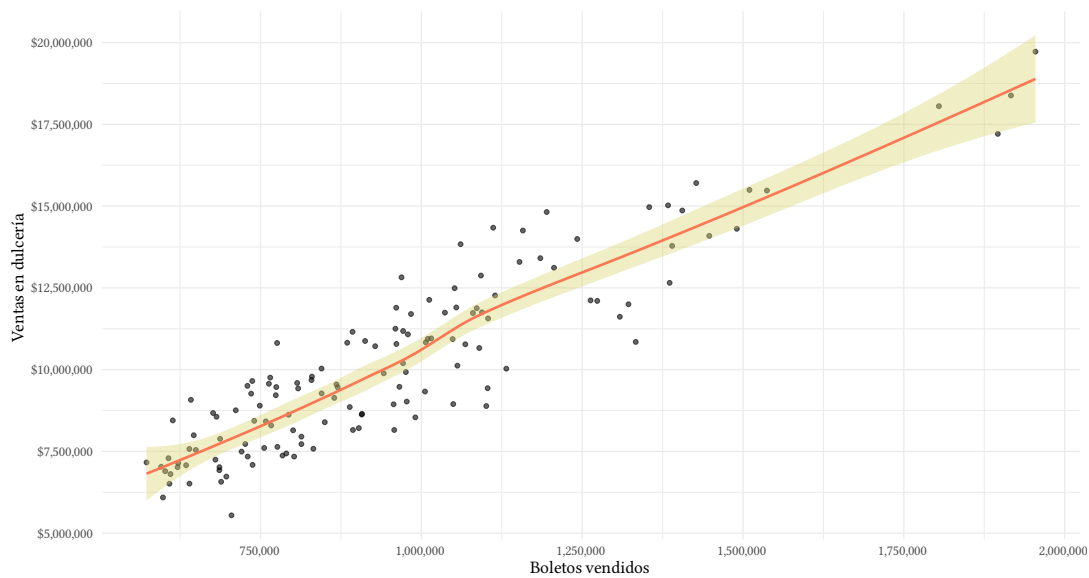


Fig. 27. Asistentes semanales promedio v. venta de palomitas

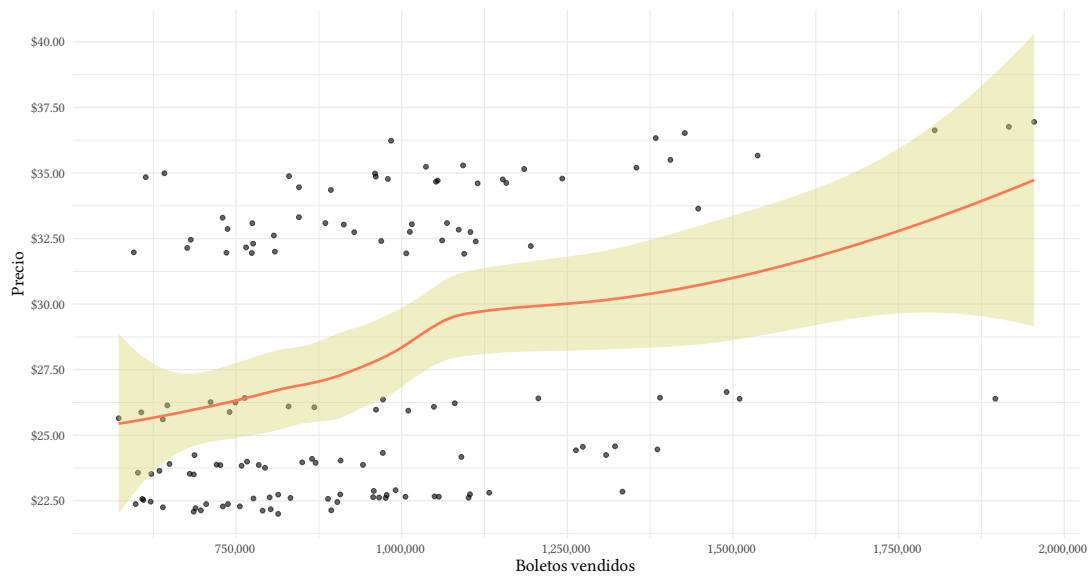


Fig. 28. Asistentes semanales promedio v. precios de bebidas embotelladas

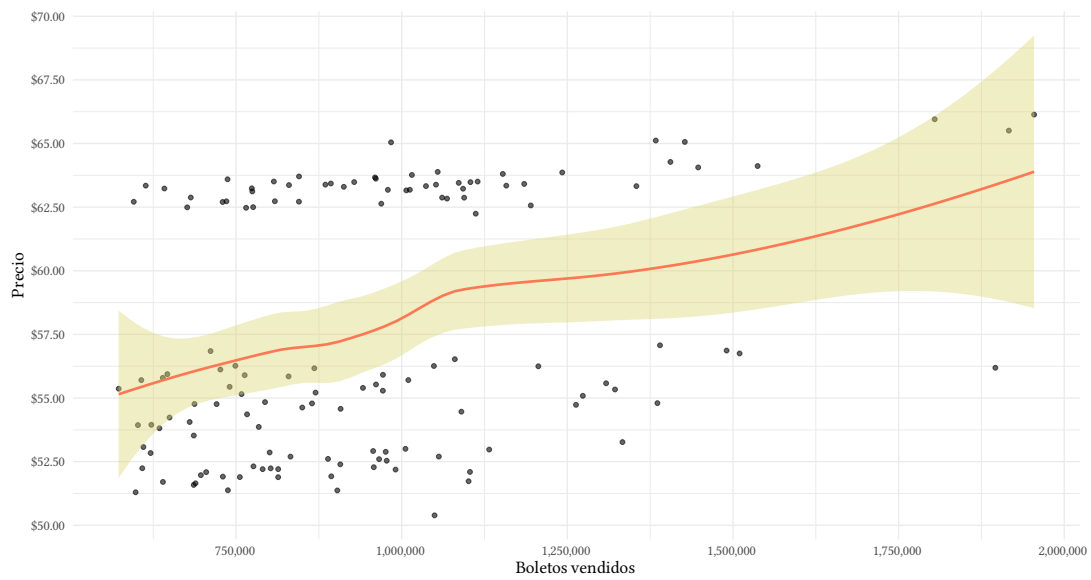


Fig. 29. Asistentes semanales promedio v. precios de bebidas de cafetería



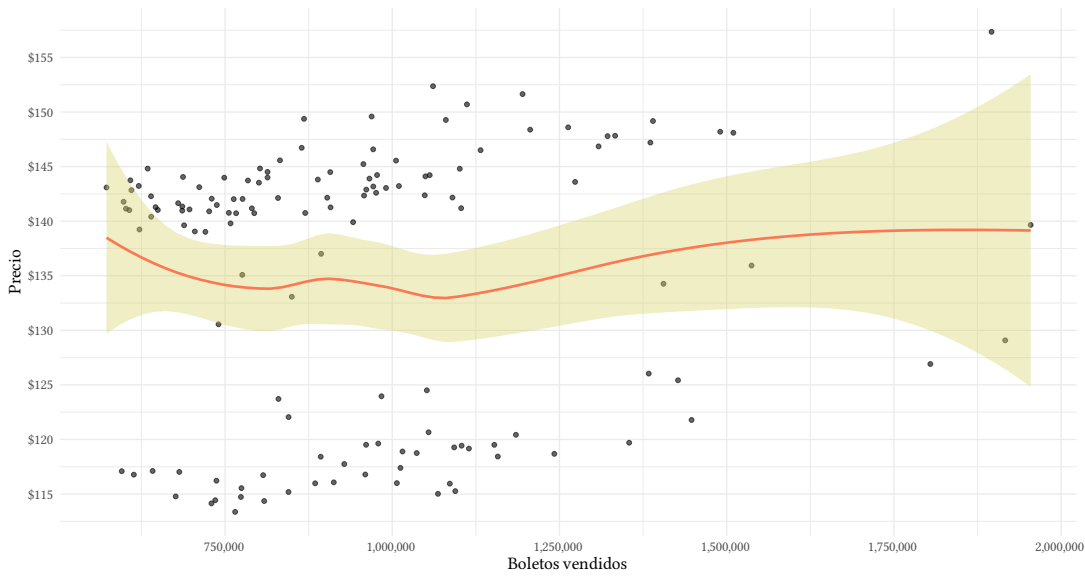


Fig. 30. Asistentes semanales promedio v. precios de combos

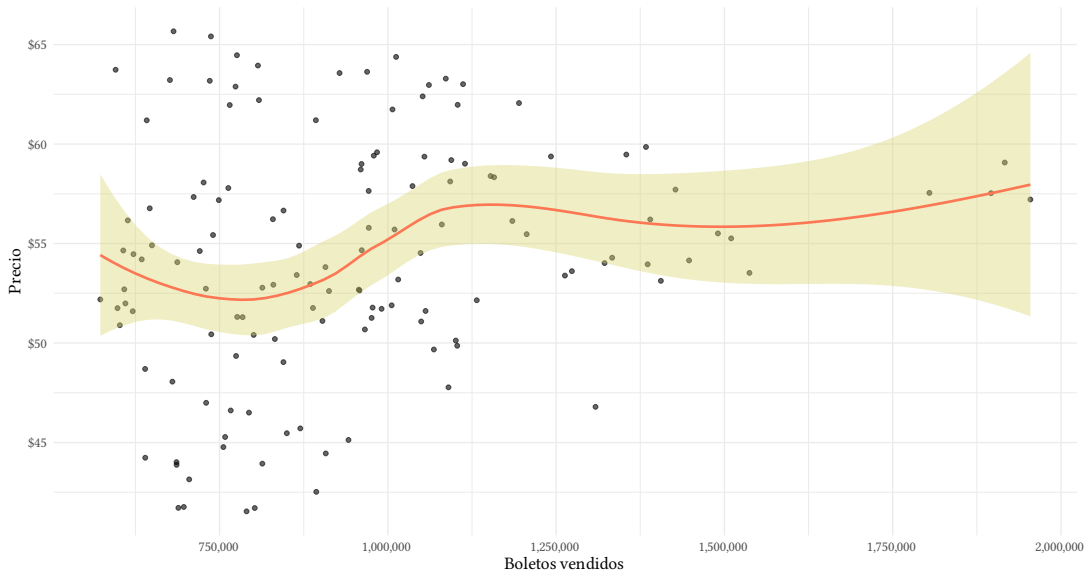


Fig. 31. Asistentes semanales promedio v. precios de postres, nachos, nuggets, etc.

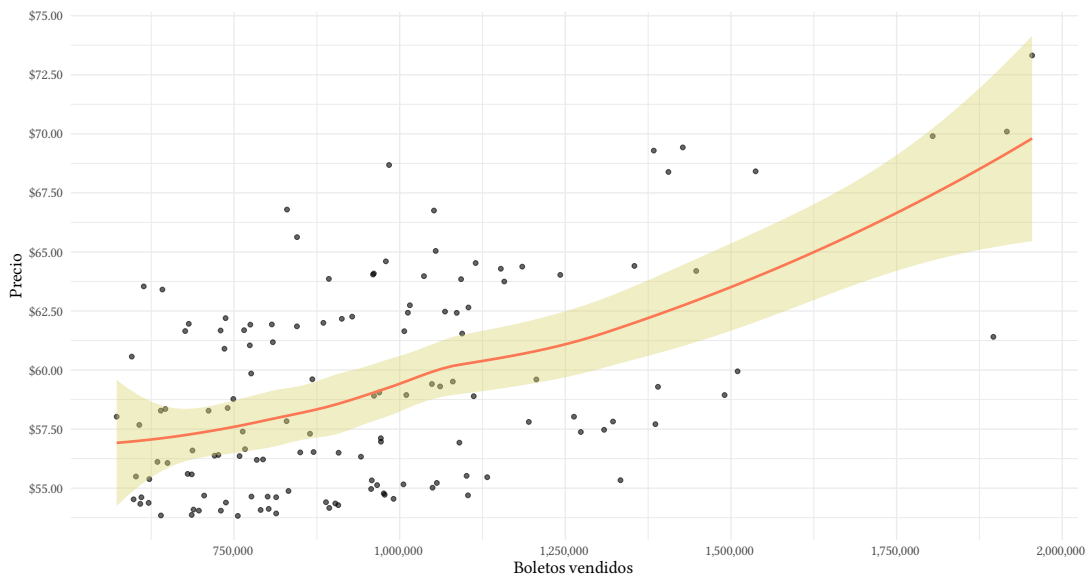


Fig. 32. Asistentes semanales promedio v. precios de palomitas

Producto	Ventas y número de asistentes	Precio y número de asistentes
Bebidas embotelladas	0.7991	0.3659
Bebidas de la cafetería	0.4715	0.3491
Combos	0.9589	0.0472
Otros	0.9445	0.2161
Palomitas	0.9106	0.5227

Cuadro 7. Correlación de Pearson entre ventas y número de asistentes y precio y número de asistentes

Recordemos que, entre el número que indica la correlación (que se encuentra entre el 0 y al 1) se acerque más al 1, más relacionado estará y al contrario con el 0; mientras más corta sea la distancia con este número, menos relación habrá entre las variables. En el caso estudiado, la correlación más significativa es la que hay entre las ventas de combos y el número de asistentes, con 0.95 de correlación, respectivamente.

4. OBTÉN LA DISTRIBUCIÓN MUESTRAL DE LAS VENTAS TOTALES MENSUALES DE DULCERÍA, POR TIPO DE PRODUCTO Y ZONA. A PARTIR DE DICHA DISTRIBUCIÓN, CALCULA LA PROBABILIDAD DE QUE EN CADA UNA DE LAS ZONAS, LA VENTA TOTAL DE COMBOS SEA MAYOR A 45 MILLONES DE USD. COMPARA LOS RESULTADOS Y COMENTA SOBRE LAS DIFERENCIAS.

Zona 1	Zona 2	Zona 3	Zona 4
0.9999	0.3747	0.8922	0.9999

Cuadro 8. Probabilidad de que las ventas de combos rebasen los 45 millones (por Zona)

A partir de los resultados, podemos observar que las zonas en las que es altamente probable que las ventas de estos productos superen los 45 millones de USD son la Zona 1 y la Zona 4, con un más de 99 por ciento de probabilidad, respectivamente. A estas zonas, le sigue la Zona número 3, con poco más de 89 por ciento de probabilidad. Muy lejos de estas cifras se encuentra la Zona 2, pues esta cuenta con poco más de 37 por ciento de probabilidad de que sus ventas por combos rebasen los 45 millones de USD. Esto se puede explicar debido a que, en promedio, los precios que ofrece esta Zona al público son mayores a los de las demás zonas, lo que podría desincentivar el consumo de los productos en general. Además, como se puede apreciar en el Cuadro 4, la Zona 3 presenta una venta de boletos significativamente menor con respecto a las otras zonas. Por ejemplo, las ventas de la Zona 3 representan tan solo el 70 por ciento de las ventas de boletos de la Zona con más cantidad de ventas, que es la Zona 1.

5. CALCULA UN INTERVALO DE CONFIANZA PARA EL NÚMERO PROMEDIO DE ASISTENTES A LA SEMANA PARA CADA UNA DE LAS ZONAS Y RESPONDE A LA PREGUNTA: ¿LOS DATOS SUGIEREN QUE HAY MAYOR ASISTENCIA EN ALGUNA DE LAS ZONAS?

Nivel de confianza	$L_n$	$U_n$
90 %	1 022 215.063	1 111 142.536
95 %	1 013 696.989	1 119 660.610
99 %	997 048.898	1 136 308.701

Cuadro 9. Intervalos de confianza para el número promedio de asistentes en la Zona 1

Nivel de confianza	$L_n$	$U_n$
90 %	719 289.389	781 203.084
95 %	713 358.877	787 133.596
99 %	701 768.029	798 724.444

Cuadro 10. Intervalos de confianza para el número promedio de asistentes en la Zona 2

Nivel de confianza	$L_n$	$U_n$
90 %	896 990.174	982 468.581
95 %	888 802.475	990 656.279
99 %	872 800.083	1 006 658.672

Cuadro 11. Intervalos de confianza para el número promedio de asistentes en la Zona 3

Nivel de confianza	$L_n$	$U_n$
90 %	1 019 807.7291	1 107 297.752
95 %	1 011 427.344	1 115 678.137
99 %	995 048.357	1 132 057.123

Cuadro 12. Intervalos de confianza para el número promedio de asistentes en la Zona 4

A través de los datos obtenidos en las tablas superiores, podemos concluir que, en efecto, hay un número mayor de asistencia en ciertas zonas, sobre todo en la Zona 1. Diferentes evidencias, además de las que brindaron el cuadro 9, concuerdan con este dato; la Zona 1 tiene una mayor venta de combos y la correlación entre ventas y número de asistentes en esta Zona, es la más fuerte (0.95).

Además, la Zona 1 posee precios más competitivos en sus productos que los de otras zonas, lo que provoca que los asistentes sean más propensos a consumir los alimentos y bebidas que ofrecen en este lugar.

6. CALCULA LA CORRELACIÓN PRIMERO DEL NÚMERO TOTAL DE BOLETOS VENDIDOS A NIVEL SEMANAL A TOTAL PAÍS CON EL TOTAL DE VENTAS EN DULCERÍA Y DESPUÉS CON LA VENTA PROMEDIO EN DULCERÍA. REPORTA LOS RESULTADOS Y ELABORA UNA EXPLICACIÓN SOBRE LOS RESULTADOS QUE ENCUENTRES.

$r_{x,y}$	Venta total de tickets
Total de ventas en dulcería	0.9731
Ventas promedio en dulcería	0.9731

Cuadro 13. Coeficientes de correlación de Pearson

Con el fin de comprender la razón por la cual los coeficientes de relación son iguales en ambos casos, es necesario conocer el funcionamiento del método de Pearson. Éste mide la dependencia lineal sobre dos estadísticos muestrales relacionados entre sí por un tercero.

Es claro que los totales de venta semanal se obtienen sumando aquellos datos correspondientes a una misma semana y, de forma análoga pero con la media, las ventas promedio en dulcería.

Sin embargo, es necesario observar que las ventas promedio representan una razón del total de ventas, pues, por definición,  $\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$ . Esto último implica la conservación del coeficiente de correlación de Pearson, puesto que la dispersión de los datos contemplaría el mismo comportamiento, únicamente a una escala diferente.

7. A PARTIR DE LA INFORMACIÓN QUE SE TIENE SOBRE LAS SEMANAS BLOCKBUSTER, OBTÉN UN INTERVALO DE CONFIANZA PARA DETERMINAR SI ES POSIBLE ASEGURAR QUE LAS SEMANAS BLOCKBUSTER VENDEN MÁS EN DULCERÍA QUE LAS SEMANAS REGULARES.

Nivel de confianza	$L_n$	$U_n$
90 %	\$ 102 346 809.121	\$ 128 706 573.313
95 %	\$ 99 821 893.048	\$ 131 231 489.386
99 %	\$ 94 887 087.811	\$ 136 166 294.623

Cuadro 14. Intervalos de confianza para el promedio de ventas en semanas *blockbuster*

Nivel de confianza	$L_n$	$U_n$
90 %	\$ 82 470 673.20	\$ 88 777 583.02
95 %	\$ 81 866 554.82	\$ 89 381 701.40
99 %	\$ 80 685 839.72	\$ 90 562 416.50

Cuadro 15. Intervalos de confianza para el promedio de ventas en semanas *no-blockbuster*

Recordemos que las semanas *blockbuster* son aquellas en las que hubo algún estreno de películas *blockbuster*, cuya estadia en cartelera generó en taquilla más de 15 millones de dólares. A partir de los resultados que arrojan los cuadros 14 y 15, podemos afirmar que, durante las semanas *blockbuster*, sí hay un mayor número de ventas en dulcería que en las semanas *no-blockbuster*. Este fenómeno tiene diversas explicaciones; en las semanas *blockbuster* hay una mayor cantidad de asistentes, probablemente porque: la película anunciada en cartelera es muy esperada; porque reconocidas actrices y actores participan en ella; porque fue premiada internacionalmente o porque es parte de un mundo cinematográfico con muchos adeptos, tales como *Marvel* o *Disney Pixar*. De igual mamnera, la correlación entre el número de asistentes y las ventas en dulcería son positivas y muy cercanas a uno. Tal es el caso de las ventas de bebidas embotelladas y el número de asistentes (0.7991), la venta de combos y el número de asistentes (0.9589), la venta de otros (nuggets, postres, etc.) y el número de asistentes (0.9445) y entre las palomitas y la variable ya mencionada (0.9106). Todos estos datos nos llevan a indicar que, entre más asistentes haya, más ventas en dulcería habrá.

8. UNO DE LOS CRITERIOS DE LA COMPAÑÍA ES QUE SI LA VENTA DE BOLETOS EN UNA SEMANA DADA ES SUPERIOR A LOS 5 MILLONES DE BOLETOS, ENTONCES SE CONSIDERA UNA SEMANA *OUTSTANDING*. POR OTRO LADO, UNO DE LOS PRINCIPALES INDICADORES PARA EVALUAR EL DESEMPEÑO DE LA COMPAÑÍA SON LOS MOMIOS (EN INGLÉS, *ODDS*) DE OBTENER UNA SEMANA *OUTSTANDING*. SE DEFINE COMO RAZÓN DE MOMIOS A LA PROPORCIÓN (PROBABILIDAD) DE VECES QUE UN SUCESO OCURRA CONTRA QUE NO OCURRA; ES DECIR:  $odds = \frac{p}{1-p}$ . OBTÉN EL ESTIMADOR MÁXIMO VEROSÍMIL DE LOS MOMIOS DE UNA SEMANA *OUTSTANDING* PARA LA COMPAÑÍA CONSIDERANDO TODA LA HISTORIA.

Es fácil ver que sólo se necesita de una condición para considerar a una semana *outstanding*, o no, por lo que nos enfrentamos a una distribución Bernoulli con parámetro  $p$  desconocido, cuyo estimador máximo verosímil es  $\hat{p} = \bar{X}$ .

Asimismo, por la *propiedad de invarianza* de los estimadores máximo verosímiles,  $\frac{\bar{X}}{1-\bar{X}}$  es el estimador máximo verosímil de  $\frac{p}{1-p}$ .

En este caso, nuestro estimador máximo verosímil es:

$$MLE = 0,1637931.$$

## 9. CONCLUSIÓN

Este análisis de los números arrojados por el desempeño de una empresa de entretenimiento estadounidense se realizó con el objetivo de dar recomendaciones claras y concisas a la misma, y, a su vez, con el fin de mejorar su funcionamiento y maximizar sus ganancias en cada una de las 4 zonas en las que opera. Con base en los datos desplegados a lo largo de este documento, es posible asegurar que ha disminuido la venta en dulcería a través de los años (periodo 2017-2019) en Florida, el Atlántico Central y el Sureste. En cambio, en Nueva Inglaterra han aumentado a tasas decrecientes.

Por otro lado, el comportamiento de ventas en dulcería no es similar en las 4 zonas; las zonas 1 y 4 son las que tienen mayores ventas en dulcería, así como un mayor número de venta de boletos, por lo que podemos concluir que una mayor audiencia implica mayores ventas en dulcería. Sin embargo, el consumo de los productos que ofrece la empresa es diferenciado; los más vendidos son los combos y *otros*, seguidos por las palomitas. Estos resultados podrían parecer ilógicos si solo consideramos como variable de decisión el precio, ya que los combos tienen el precio promedio más caro en el menú (\$132.77), a diferencia de las bebidas embotelladas, por ejemplo (\$28.10). Sin embargo, la relación entre el precio y el número de asistentes, también es diferente. En el caso de los combos, la correlación de Pearson arroja 0.0472, lo que indica que es casi independiente el valor de los combos y el número de asistentes en el lugar. Por el contrario, existe una correlación importante, más no determinante, en el caso de las palomitas (0.5227).

Nuestra recomendación es dar más promoción tanto a los combos, como a las palomitas y también a los productos que caen en la categoría *otros*, pues son los más vendidos. No recomendamos disminuir el precio de los combos, pues no tendrá un efecto significativo en las ventas. Por el contrario, recomendamos disminuir el precio de las palomitas, pues esto sí podría aumentar la cantidad demandada de estas. Asimismo, como las ventas de bebidas embotelladas individuales no tienen mucho éxito, se podrían incluir en la estructura de los combos para aumentar su consumo (sin quitar la venta individual de estos productos, que también trae ganancias). Por último, exhortamos a la empresa a hacer más atractiva la sucursal en la Zona 2, a través del estreno de más *blockbusters* o la proyección exclusiva de festivales anuales de cine, como el *Festival de Cine de Morelia*, en México, para atraer más clientes.