

Análisis Exploratorio de Datos

Sabor en números:

Análisis de restaurantes en España

Proyecto: Comprensión de Datos

Ciencia de Datos - UPV

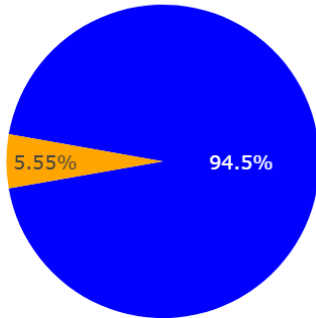
PABLO LUCAS MORA, KENTO KAMAKURA GIMENO, MIGUEL
GIL JIMENEZ, DANIEL HERNANDO GRABOLEDA

ÍNDICE

ANÁLISIS UNIVARIANTE	2
MICHELÍN.....	2
RESERVABLE.....	2
PRECIO MEDIO	3
RESERVAS ÚLTIMA SEMANA	3
RESEÑA DE COMIDA	4
ANÁLISIS MULTIVARIANTE	5
SALARIO MEDIO POR COMUNIDAD	5
PRECIO MEDIO DEL RESTAURANTE POR COMUNIDAD	5
TIPO DE COMIDA SEGÚN LA COMUNIDAD	6
CANTIDAD DE RESTAURANTES POR POBLACIÓN POR PROVINCIA.....	6
PRECIO MEDIO POR TIPO DE COMIDA	7
CORRELACIONES.....	8
PRECIO MEDIO Y SALARIO MEDIO	9
DISTINCIÓN Y VALORACIÓN DE COMIDA.....	9
PRECIO MEDIO Y VALORACIÓN DE COMIDA	9
NÚMERO DE PREMIOS Y PRECIO MEDIO	9
RESERVAS EN LA ÚLTIMA SEMANA CON RESTAURANTES MICHELÍN	9
PRECIO MEDIO Y RESTAURANTES MICHELÍN	9
SERVICIO Y AMBIENTE.....	10
RESERVAS EN UNA SEMANA Y VALORACIÓN DE COMIDA	10
NÚMERO DE FOTOS, DE COMENTARIOS Y DE VALORACIONES	10
RESTAURANTES POR POBLACIÓN Y SALARIO MEDIO ANUAL.....	10
CONCLUSIONES AED	11

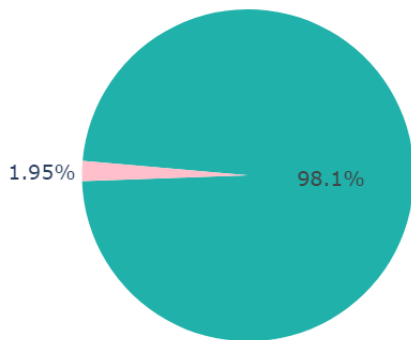
ANÁLISIS UNIVARIANTE

MICHELÍN



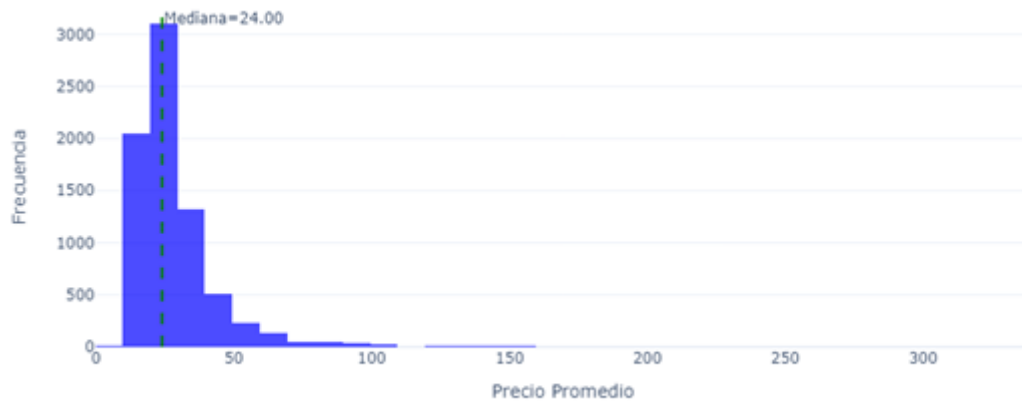
Como vemos, solo el 5'55% de los restaurantes tiene una o más estrellas michelín. Esto demuestra en gran parte la dificultad que conlleva conseguir una estrella michelín en España

RESERVABLE



Aquí se muestra que actualmente en España el servicio de reserva es prácticamente del 98%, resaltando que a la gran mayoría de los restaurantes se puede acudir mediante la reserva previa de la mesa.

PRECIO MEDIO



En primer lugar, eliminamos las columnas cuyo precio medio era menor a 6 euros y es que por muy barato que sea el restaurante, no es lógico que en España puedas comer por 6 euros. El precio medio por comensal según el restaurante es otra variable muy dispersa y tiene un rango de 307 valores, indicando la presencia de muchos valores distintos y un comportamiento variado como vemos en el diagrama de barras. Por ello decimos que presenta una asimetría positiva debido a una mayor concentración de valores al principio y la presencia de valores atípicos como 307, que hacen que la medida adecuada de centralización sea la mediana, cuyo valor, 24, es mucho más razonable. Por último, es muy reseñable la curtosis de 632, que es mucho mayor a 0, evidenciando una forma muy leptocúrtica.

RESERVAS ÚLTIMA SEMANA



Esta es de las variables más dispersas y tiene un coeficiente de variación del 229%, junto con un rango de 1133 valores, indicando la presencia de muchos valores distintos y un comportamiento irregular como el diagrama de barras. Es por ello por lo que presenta una asimetría positiva debido a una mayor concentración de valores al principio y la presencia de valores atípicos como 1133, que hacen que la medida adecuada de centralización sea la mediana o media recortada. Por último, es muy reseñable la curtosis de 1014, que es mucho mayor a 0, evidenciando una forma muy leptocúrtica.

RESEÑA DE COMIDA

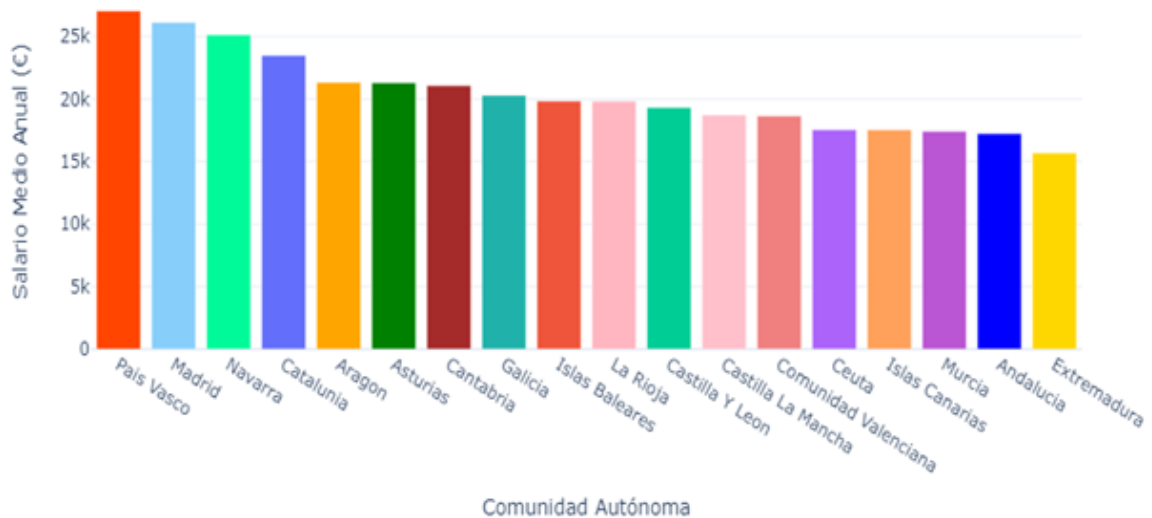


En la valoración de la comida, observamos que el rango de valores no es extenso y que todos se concentran entre el 4 y el 10, siendo los números más elevados los con mayor presencia de valores, así que presenta una asimetría negativa por una mayor concentración de valores al final y una menor al principio.

ANÁLISIS MULTIVARIANTE

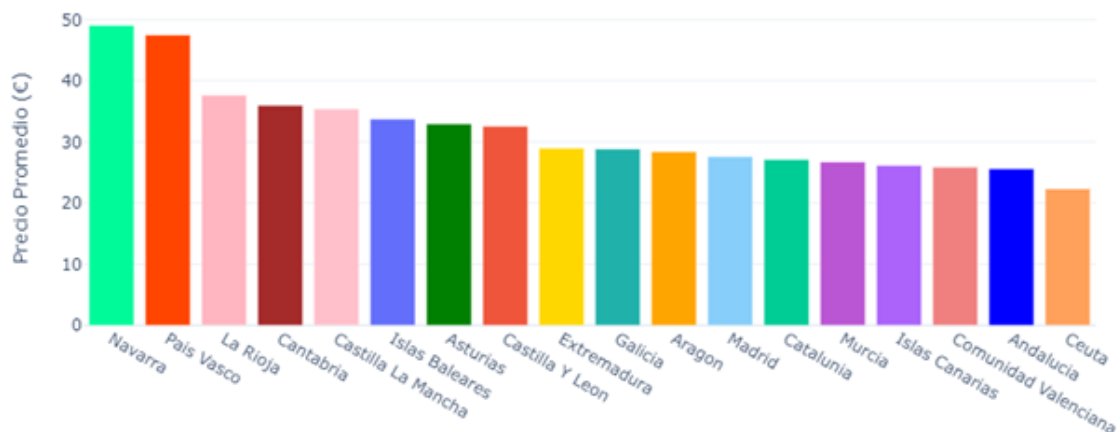
A continuación, mostramos las siguientes gráficas que recogen los salarios y precios medios según la ubicación, que nos guiarán a la hora de sacar conclusiones

SALARIO MEDIO POR COMUNIDAD



En primer lugar, observamos el sueldo medio de los españoles según la comunidad autónoma, liderado por País Vasco y Madrid y con Extremadura como comunidad con el salario más bajo.

PRECIO MEDIO DEL RESTAURANTE POR COMUNIDAD



La siguiente gráfica muestra por orden el precio medio por persona de una comida por

comunidad. Esto nos revela que la comunidad más cara es Navarra, con un precio medio de casi 50€, frente a Ceuta, donde apenas sobrepasa los 20€.

TIPO DE COMIDA SEGÚN LA COMUNIDAD

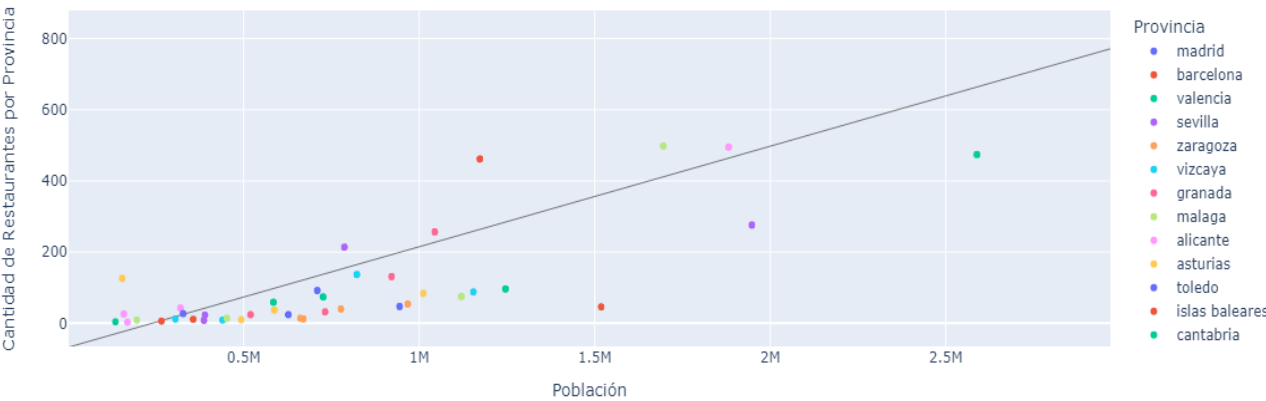
COMUNIDAD AUTÓNOMA	COMIDA MÁS CONSUMIDA 1	COMIDA MÁS CONSUMIDA 2	COMIDA MENOS CONSUMIDA 1	COMIDA MENOS CONSUMIDA 2
Andalucía				
Galicia				
C. Valenciana				
Madrid				
Cataluña				
Islas Canarias				
Asturias				

Hemos recogido una tabla que muestra los tipos de comida más repetidos en las siguientes comunidades:

Se puede concluir como en las grandes ciudades aparece más tipo de comida internacional, debido al tráfico de turistas y extranjeros, siendo la comida española y sobre todo la mediterránea la principal. Por otro lado, en las comunidades más pequeñas y de menor población predomina la comida local. A pesar de ello, las comunidades con fuerte identidad gastronómica suelen superar al resto (véase C. Valenciana, Cataluña, Andalucía...)

No obstante, el tipo de comida menos consumido en las diferentes comunidades autónomas son aquellos alimentos propios de la cocina internacional menos representada, como la comida coreana, india, venezolana, peruana...

CANTIDAD DE RESTAURANTES POR POBLACIÓN POR PROVINCIA



Este gráfico de puntos recoge los datos de la cantidad de restaurantes por provincias. La línea de tendencia muestra que estas variables siguen una relación proporcional y totalmente lineal, es decir, cuanta más población hay en una provincia, más cantidad de restaurantes posee

PRECIO MEDIO POR TIPO DE COMIDA

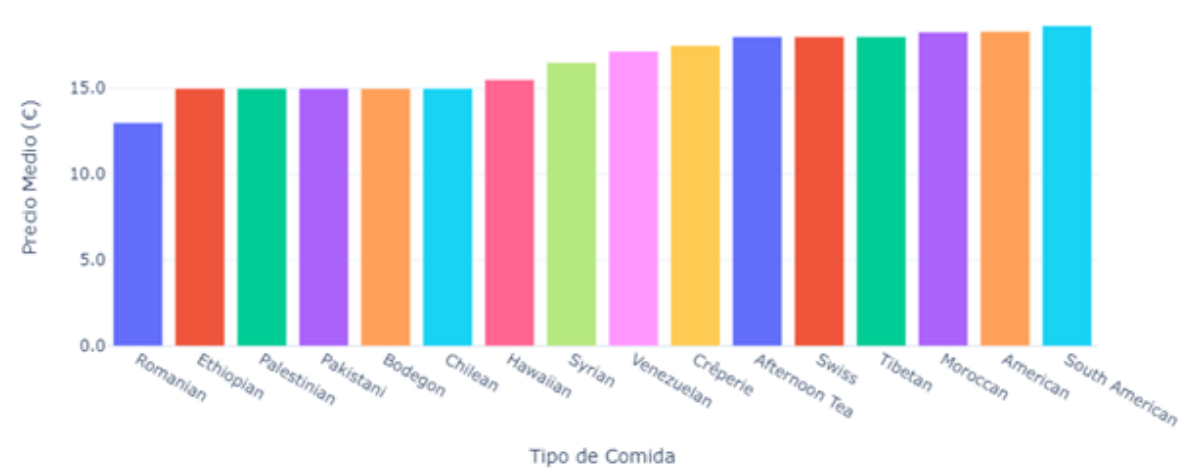
En los siguientes gráficos de barras hemos recogido los principales tipos de comida y los hemos representado según su precio medio:

- **Mayor a menor precio**



Como vemos, la comida local y del norte de España, sobre todo del País Vasco, es la más cara.

- **Menor precio**



Mientras que la comida proveniente de países como Etiopía, Palestina, Pakistán y Rumanía es la más barata.

CORRELACIONES

Aquí mostramos las diferentes correlaciones de Pearson estudiadas en nuestro proyecto mediante la siguiente tabla, evitando así poner el gráfico de correlaciones de Pearson en cada una de las relaciones, y es que sólo nos interesa el número, por lo que simplificamos el resultado de esta forma:

Variables estudiadas	Correlación de Pearson
Precio medio – salario medio	0,03
Distinción - Valoración de comida	0,83
Precio medio - Valoración de comida*	0,22
Número de premios - Precio medio	0,55
Reservas en una semana - Restaurantes Michelin*	0,07
Estrella Michelin - Precio medio	0,59
Servicio - Ambiente	0,79
Reservas en una semana - Valoración de comida	0,09
Número de comentarios - Número de fotos	0,1
Número de fotos - Valoraciones	0,1
Número de comentarios - Valoraciones	0,98
Población - Cantidad de restaurantes	0,91
Salario medio - Cantidad de restaurantes	0,32
Salario medio - Población	0,43

A continuación, explicamos de forma más detallada el significado del número obtenido en cada relación:

PRECIO MEDIO Y SALARIO MEDIO

Se observa que la relación precio medio y salario medio es prácticamente nula y puede estar influenciada por una variedad de factores adicionales, como la diversidad económica, la demanda y la oferta, los costos operativos y la segmentación del mercado.

DISTINCIÓN Y VALORACIÓN DE COMIDA

La distinción de un restaurante y la valoración de su comida están estrechamente relacionadas, ya que los estándares de calidad, innovación, ingredientes y consistencia en la oferta gastronómica son elementos clave que influyen tanto en la obtención de reconocimientos como en la satisfacción de los comensales.

-

PRECIO MEDIO Y VALORACIÓN DE COMIDA

Llama también la atención la escasa relación entre el precio medio y la valoración, pues se tiende a asociar una buena valoración con mayor precio del restaurante. Este gráfico de correlaciones Pearson demuestra que no hay mucha conexión, por lo que mayor precio medio no implica una mejor comida y viceversa.

NÚMERO DE PREMIOS Y PRECIO MEDIO

En cuanto a las variables Numero_awards (número de premios del sitio) y Precio Medio, observamos una ligera conexión, por lo que concluimos que hay cierta relación entre los precios sobre los que suele rondar un restaurante para que obtenga premios.

RESERVAS EN LA ÚLTIMA SEMANA CON RESTAURANTES MICHELÍN

Esta se trata de la correlación que más se salió de nuestras expectativas al realizar el análisis multivariado de correlaciones, dado que los restaurantes con estrella Michelin son muy populares por su calidad y reciben un muy elevado número de reservas por semana. Una vez nos pusimos a investigar a qué se debe esto, llegamos a la conclusión de que las reservas en este tipo de restaurantes se realizan a través de sus propias páginas web, y no mediante The Fork. Esto explicaría por qué en nuestra base de datos una gran cantidad de estos restaurantes no reciben reservas.

PRECIO MEDIO Y RESTAURANTES MICHELÍN

Como era de esperar, comer en un restaurante con estrellas Michelin suele ser considerablemente más caro que en otros tipos de restaurantes. Los comensales pagan

no solo por la comida, sino por la experiencia culinaria completa, que incluye servicio, ambiente y la innovación en cada plato.

SERVICIO Y AMBIENTE

El servicio y el ambiente de un restaurante están fuertemente relacionados y se complementan entre sí para crear una experiencia gastronómica más satisfactoria para los comensales. Un ambiente acogedor y bien diseñado puede realzar la calidad del servicio, mientras que un servicio amable y atento puede mejorar la percepción del ambiente por parte de los clientes.

RESERVAS EN UNA SEMANA Y VALORACIÓN DE COMIDA

Para nuestra sorpresa, el número de reservas en la última semana no tiene que ver con la valoración de comida. Esto quiere decir que, porque un restaurante esté muy reservado o frecuentado, no está comprometida su calidad.

NÚMERO DE FOTOS, DE COMENTARIOS Y DE VALORACIONES

En este gráfico analizamos tres variables: número de valoraciones, de comentarios y la cantidad de fotos subidas del restaurante.

A pesar de no hallar relación entre el número de fotos y las valoraciones y/o comentarios, el número de comentarios y valoraciones sí que presenta una relación lineal, pues los clientes que dejan un comentario dejan seguidamente una puntuación sobre el sitio donde han consumido.

RESTAURANTES POR POBLACIÓN Y SALARIO MEDIO ANUAL

En el siguiente gráfico observamos las correlaciones que existen entre la cantidad de restaurantes por cada comunidad, la población de esta, y su salario medio anual. Al observar los valores que nos da, destaca la fuerte relación que existe entre la población y el número de restaurantes, cosa que si nos paramos a pensar tiene mucho sentido ya que cuanto mayor número de habitantes, más clientes y necesidades de comida, y por tanto más restaurantes. Esta relación, además de por la correlación de Pearson, también la evidenciamos anteriormente mediante un gráfico de puntos en el que vimos la proporcionalidad entre los restaurantes y la población. Y por otro lado vemos algo de relación entre la población y el salario medio, y el salario medio y la cantidad de restaurantes, pero al no superar el 0,5 no consideramos que sea un valor representativo para definir sus relaciones.

CONCLUSIONES AED

El objetivo básico planteado al inicio del trabajo fue comprobar qué es lo que influye en el precio de un restaurante en España y saber si influyen ciertos factores como la localización, la renta per cápita, el tipo de comida, etc. Para ello, se han analizado los resultados de 7624 restaurantes de las diferentes provincias españolas y se han establecido diferencias y relaciones entre ellas que nos han ayudado a tomar ciertas conclusiones y de algún modo, 'solucionar' el problema.

En primer lugar, consultamos los datos de las comunidades autónomas con mayor salario medio anual, y mediante el uso de gráficas hacernos una mejor idea de dónde se sitúa cada comunidad de España, observando que el sueldo medio de los españoles está liderado por País Vasco y Madrid y con Extremadura como comunidad con el salario más bajo.

De la misma forma, consultamos en nuestra base de datos los tipos de comida según su precio y los ordenamos para facilitar nuestras conclusiones en el futuro.

Con la variable precio medio, un parámetro fundamental a la hora de elegir un restaurante, encontramos diversas relaciones. En primer lugar, con el salario medio, encontrando una escasa relación, influenciada por una variedad de factores adicionales. Continuando con la variable comunidad: la comunidad más cara es Navarra y el País Vasco, con un precio medio de casi 50€, frente a Ceuta, donde apenas sobrepasa los 20€. Esto tiene mucho que ver con el precio medio según el tipo de comida, y es que la comida local y del norte de España (sobre todo del País Vasco) es la más cara. También es destacable la escasa relación que tiene con la valoración, pues se tiende a asociar una buena valoración con un mayor precio del restaurante, esto se podría explicar debido a que la gente no solo busca mucha calidad (altos precios), sino que busca una buena relación calidad precio, es decir, poder permitirse una buena comida a un precio razonable y asequible. Sí tiene más relación con el número de premios del restaurante, y es que un restaurante con muchos premios suele ser más caro. Lo mismo podemos concluir con la variable estrella Michelin, indicando que, como norma general, si un restaurante tiene estrella Michelin significa que el precio de éste será mayor.

En cuanto a la valoración de la comida, este se trata de otro aspecto básico a la hora de elegir un restaurante, pues ofrece una visión importante sobre uno de los aspectos críticos de la experiencia en un restaurante y puede influir significativamente en la percepción general del establecimiento por parte de los clientes potenciales. Por un lado, al relacionarlo con la distinción, encontramos una correlación entre ambas. No obstante, para nuestra sorpresa, el número de reservas en la última semana no tiene que ver con la valoración de comida. Esto quiere decir que, porque un restaurante esté muy reservado o frecuentado, no se encuentra comprometida su calidad.

Respecto al número de comentarios en un restaurante, esta es una variable que nos puede indicar su popularidad, relevancia, calidad percibida, credibilidad, etc, y sí que presenta una relación lineal con el número de valoraciones, pues los clientes que dejan un comentario dejan seguidamente una puntuación sobre el sitio donde han consumido. No obstante, el número de comentarios no implica un mayor número de fotos en la web, y es que la relación es prácticamente nula, como ya demostramos en su correspondiente apartado.

Por otro lado, también se evidencia la conexión entre las valoraciones del servicio y el ambiente, un ambiente acogedor y bien diseñado puede realzar la calidad del servicio, mientras que un servicio amable y atento puede mejorar la percepción del ambiente por parte de los clientes, lo que demuestra la importancia del ambiente que te rodea y el sentirse a gusto en un restaurante.

Según el tipo de comida, en relación con la comunidad, se puede concluir como en las grandes ciudades aparece más tipo de comida internacional, debido al tráfico de turistas y extranjeros, siendo la comida española y sobre todo la mediterránea la principal. Por otro lado, en las comunidades más pequeñas y de menor población predomina la comida local. A pesar de ello, las comunidades con fuerte identidad gastronómica suelen superar al resto. Por otro lado, el tipo de comida menos consumido en las diferentes comunidades autónomas son aquellos alimentos propios de la cocina internacional menos representada, como la comida coreana, india, venezolana, peruana...

Por último, contemplamos las correlaciones que existen entre la cantidad de restaurantes por cada comunidad, la población de esta, y su salario medio anual. Al observar los valores que nos da, destaca la fuerte relación que existe entre la población y el número de restaurantes, algo que tiene mucho sentido, ya que cuanto mayor sea el número de habitantes, más clientes y necesidades de comida, y por tanto más restaurantes. Finalmente, vemos algo de relación entre la población, el salario medio y la cantidad de restaurantes, pero al no superar el 0,5 no consideramos que sea un valor representativo para definir sus relaciones.