

MAXIMA CONCENTRACION DE RESTAURANTES POR ESTADO

Estado 1



Cant Restaurantes_1

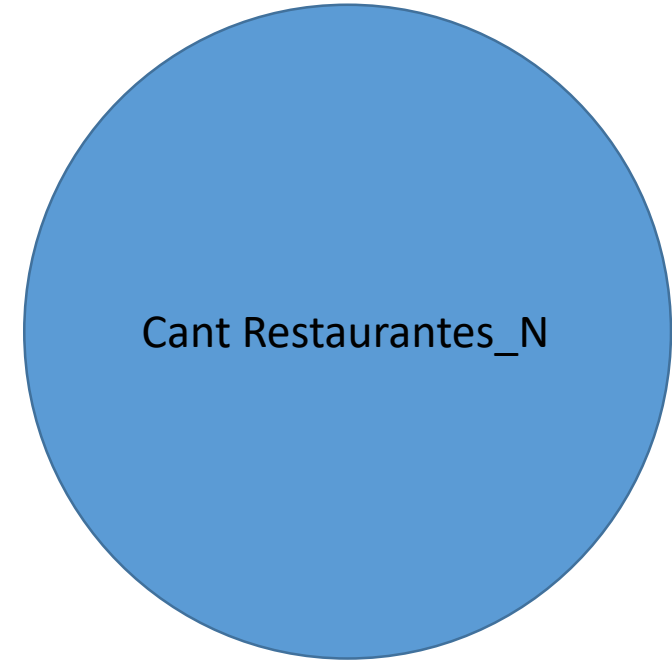
Estado 2



Cant Restaurantes_2



Estado N

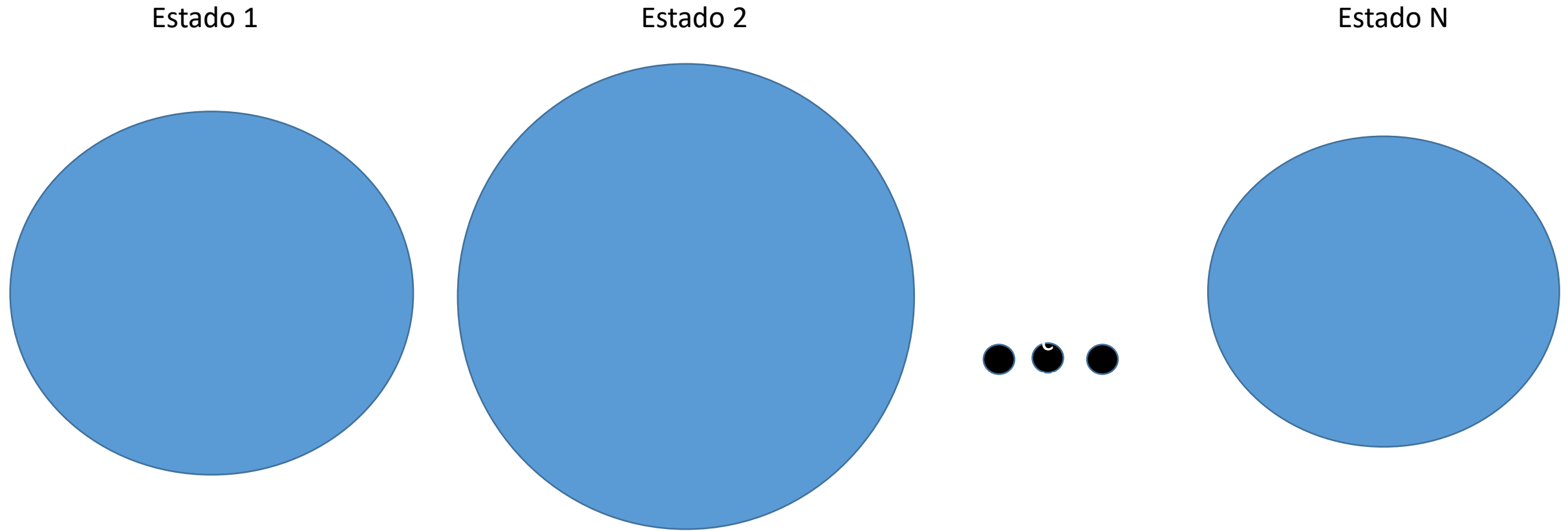


Cant Restaurantes_N

Efectivamente podemos obtener la cantidad de restaurantes por estados y de esta manera identificar cual es el estado con mayor restaurantes.

Pero esta manera tan simple de calcular un indicador podría ser injusta al momento de comparar en tres estados, pues podría haber un estado mucho mas grande (territorialmente) que el resto y por tanto contener mas restaurantes que ningún otro, y sin embargo, dicho cumulo de restaurantes están muy alejados uno de los otros, es decir, no están concentrados.

MAXIMA CONCENTRACION DE RESTAURANTES POR ESTADO



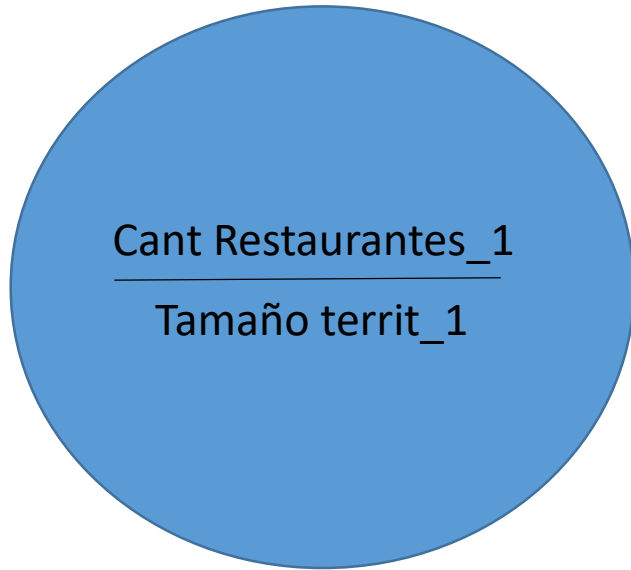
Esto representa un poco la realidad, hay estados mas grandes que otros en lo territorial y por tanto podrían haber mas comercios que en otros. Esto no necesariamente nos habla de concentración. En tal sentido, extraer la métrica mas grande cuando lo comparamos entre estados podría generar información no legitima.

$MAX (\text{Cant_rest_1} , \text{Cant_rest_1} , \text{Cant_rest_1})$

Que pasa cuando
comparamos entre
estados?

MAXIMA CONCENTRACION DE RESTAURANTES POR ESTADO

Estado 1

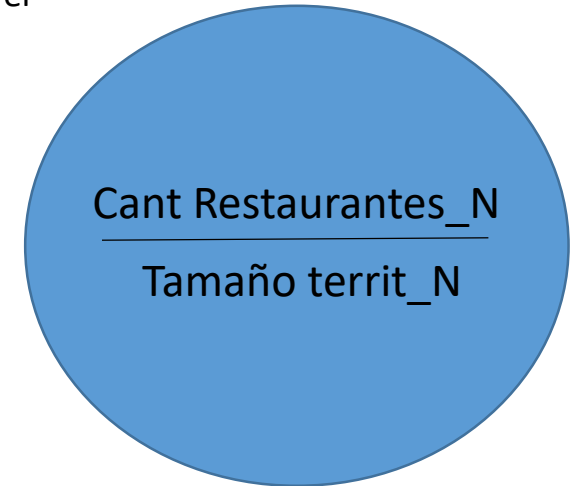


Estado 2



De esta manera si estamos ajustando el numero de restaurantes de cada estado con un factor particular y estable del mismo estado, el tamaño territorial.

Estado N



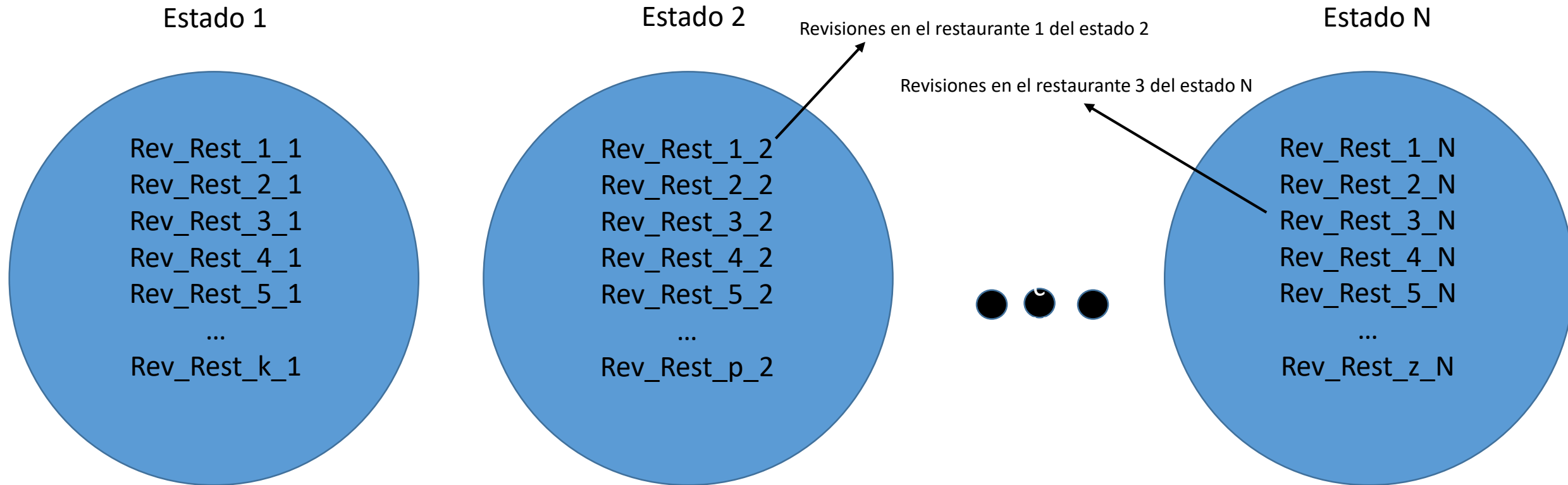
$$\text{CRE} = \frac{\text{Cant. Restaurantes}}{\text{Tamaño ter. Estado}}$$

$$\text{MAX} \left(\frac{\text{Cant Restaurantes_1}}{\text{Tamaño territ_1}}, \frac{\text{Cant Restaurantes_2}}{\text{Tamaño territ_2}}, \frac{\text{Cant Restaurantes_N}}{\text{Tamaño territ_N}} \right)$$

Es un KPI ?

Para que sirve ?
a quien le sirve ?

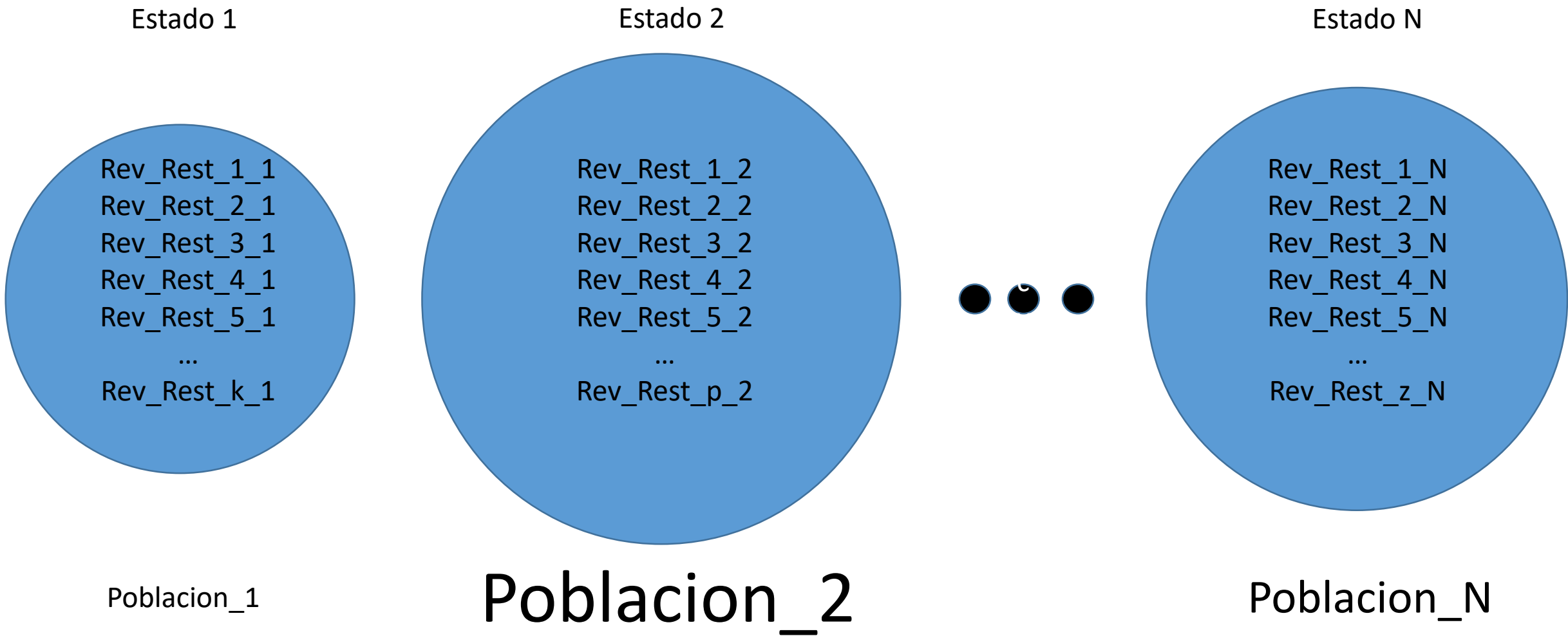
MAXIMO NIVEL DE CONSUMO



El nivel de consumo, o nivel de ventas, es una métrica que estaremos utilizando a partir de la inferencia, es decir, a partir de presumir que la participación de un consumidor en el registro de opiniones es directamente proporcional a la cantidad de personas que asisten, es decir, Mas consumen, mas participan, menos consumen, menos participan.

No importan que no tengamos los valores nominales de tales consumos, afluencia o ventas, Lo importante es utilizar esta relación.

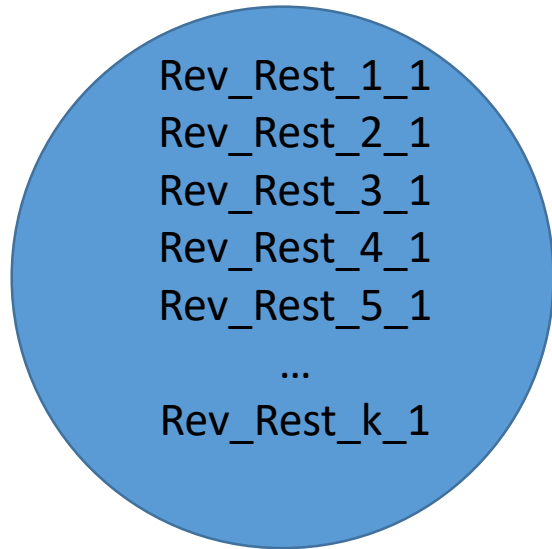
MAXIMO NIVEL DE CONSUMO



Obviamente, en lo que respecta a cualquier análisis de consumo, el factor de “la población” es fundamental. Pues hay mayor potencial de consumo en un estado en el cual la población se concentra mas. En este respecto nos gustara saber como se concentran las persona en función de la cantidad de restaurantes en el estado.

MAXIMO NIVEL DE CONSUMO

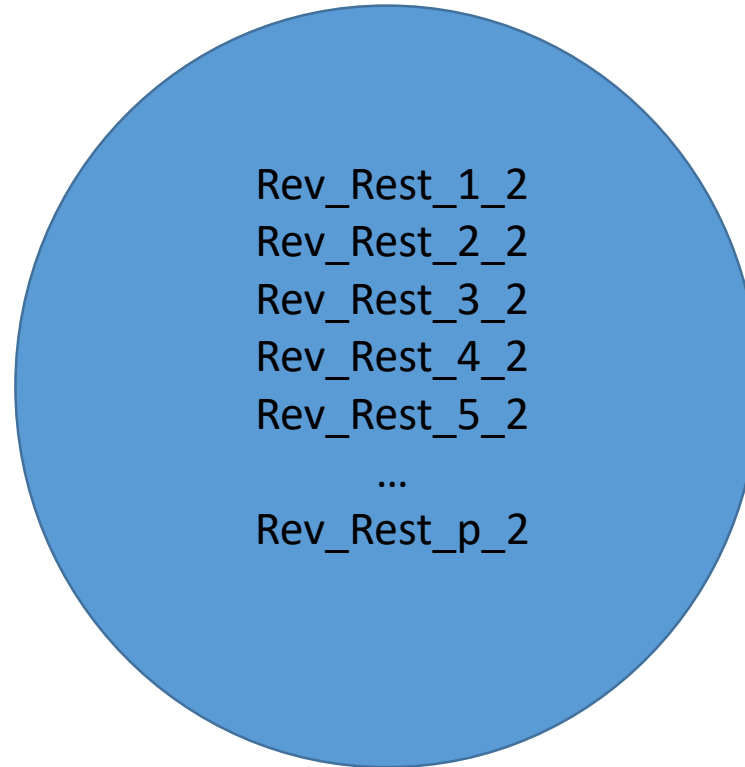
Estado 1



DCP_1

$$\frac{\text{Poblacion_1}}{\text{Can_rest_1}}$$

Estado 2

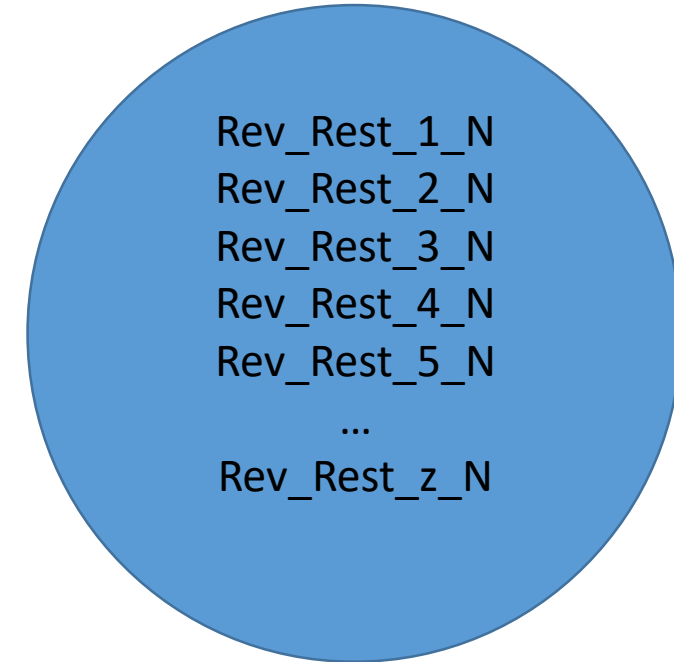


DCP_2

$$\frac{\text{Poblacion_2}}{\text{Cant_rest_2}}$$



Estado N



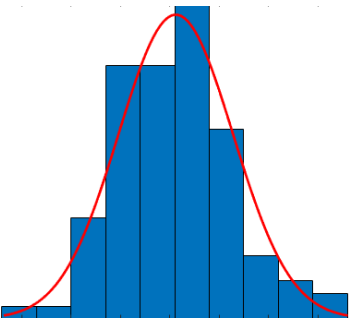
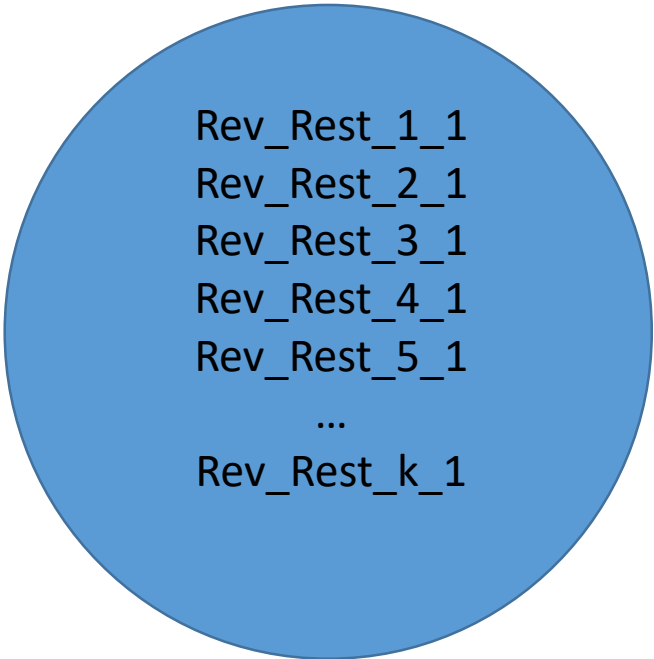
DCP_N

$$\frac{\text{Poblacion_N}}{\text{Cant_rest_N}}$$

Cuando dividimos la población entre la cantidad de restaurantes, estamos obteniendo un valor asociado a la potencialidad del consumo, lo cual si es comparable entre los estados.

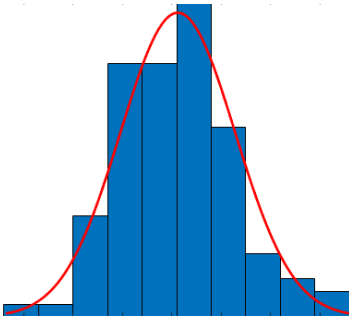
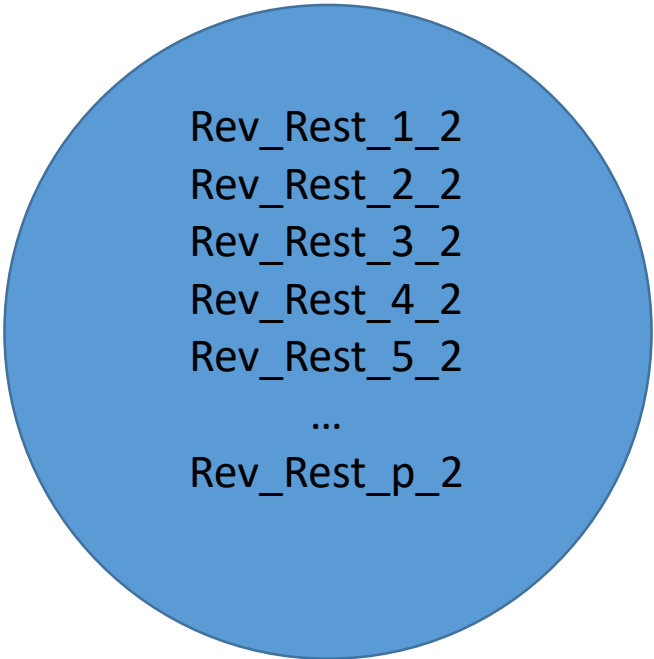
MAXIMO NIVEL DE CONSUMO

Estado 1



Min Max

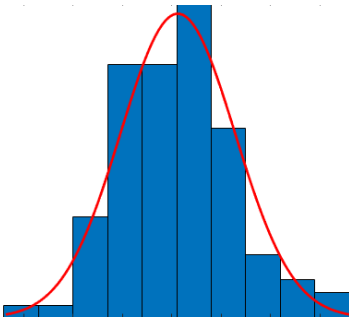
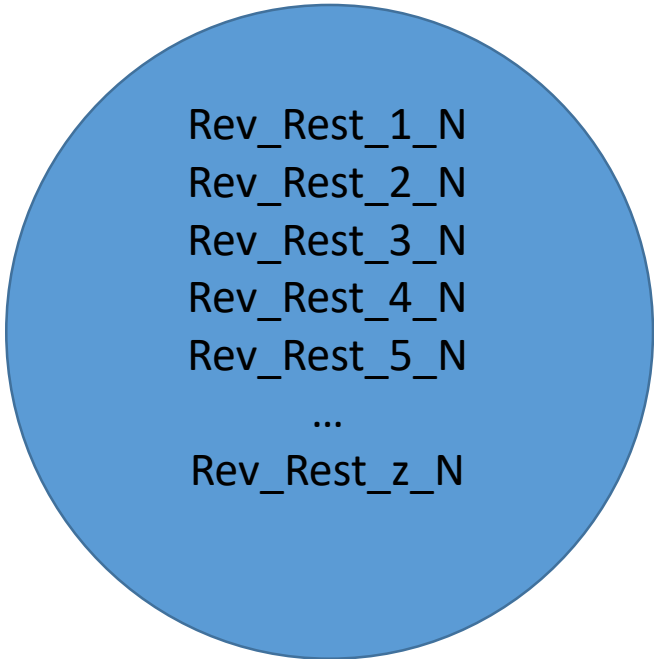
Estado 2



Min Max



Estado N

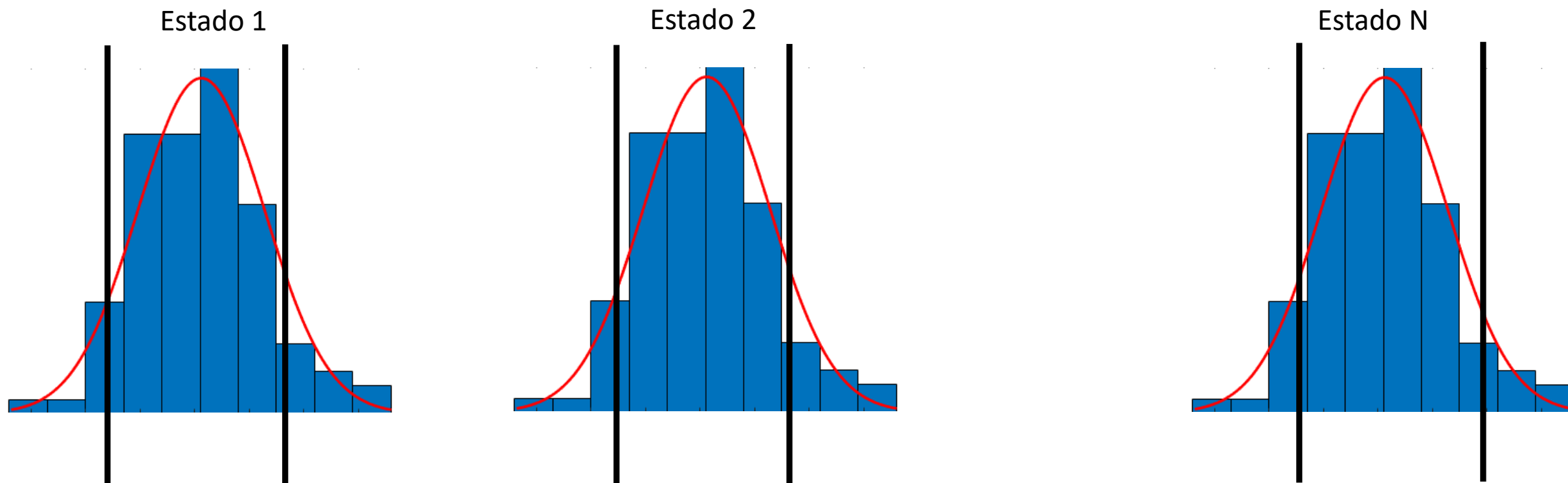


Min Max

Debemos descubrir como es el comportamiento de “las revisiones” que tiene cada comercio en su historial.

MAXIMO NIVEL DE CONSUMO

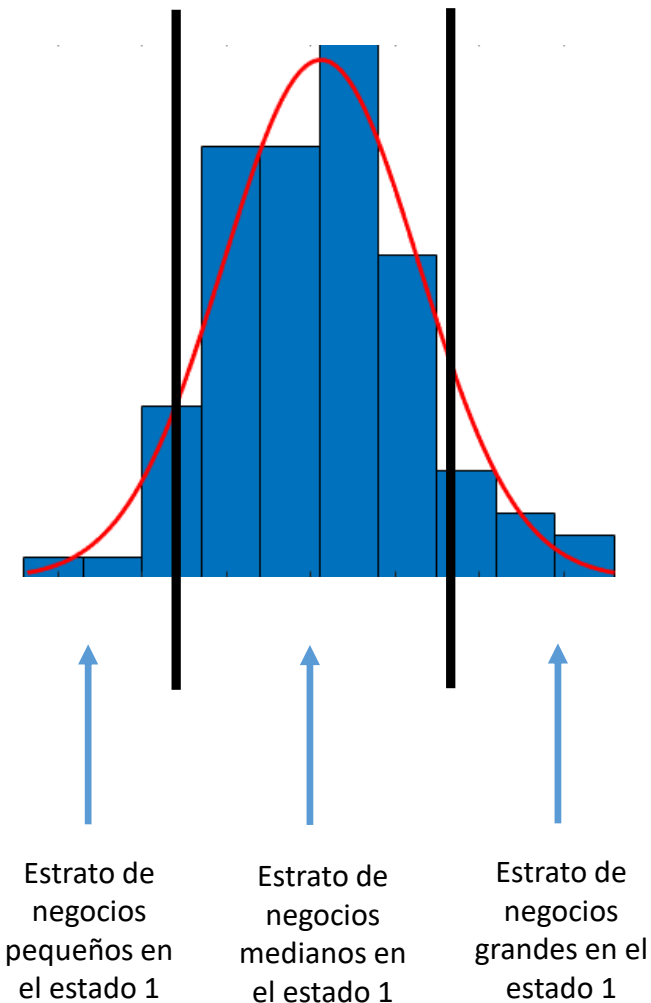
Una vez que tenemos identificado el comportamiento de frecuencia del valor “revisiones” de cada comercio por estado. Podremos entonces poder seccionar dicha distribución en dos o tres partes, con lo cual poder estratificar o segmentar, los comercios de cada estados en, pequeños, medianos y grandes. Esto desde el punto de vista del volumen de consumo, o ventas.



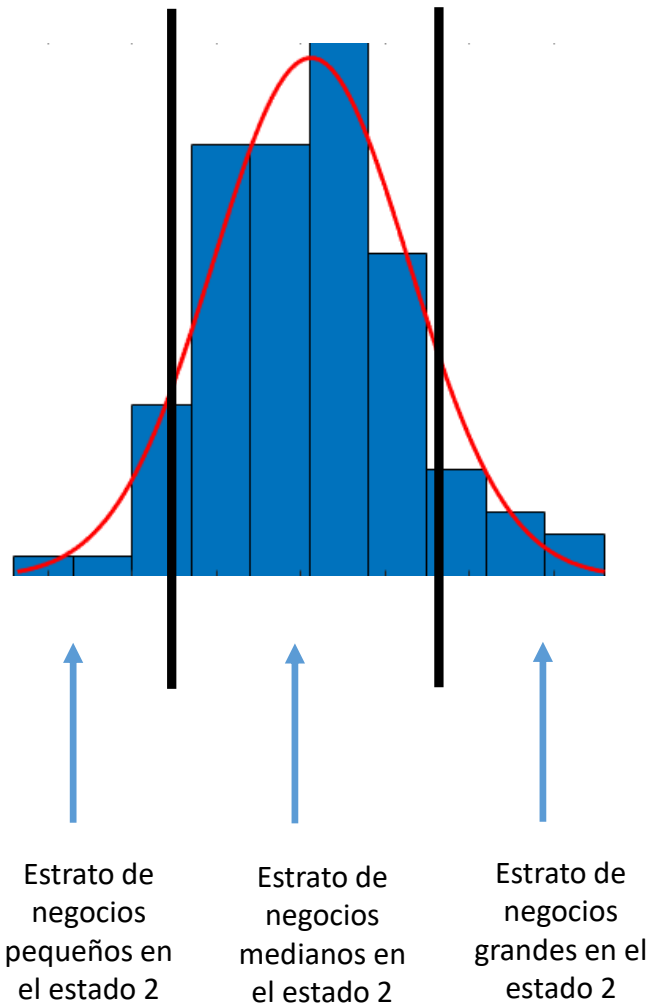
Dicho seccionamiento podremos hacerlo en función de lo que identifiquemos, sin embargo a priori un criterio básico teórico sería el de 15%, 70% y 15%, suponiendo que los resultados que obtendremos serán campanas de Gauss. Sin embargo, el contexto nos permitirá entender el seccionamiento mas conveniente en su momento.

MAXIMO NIVEL DE CONSUMO

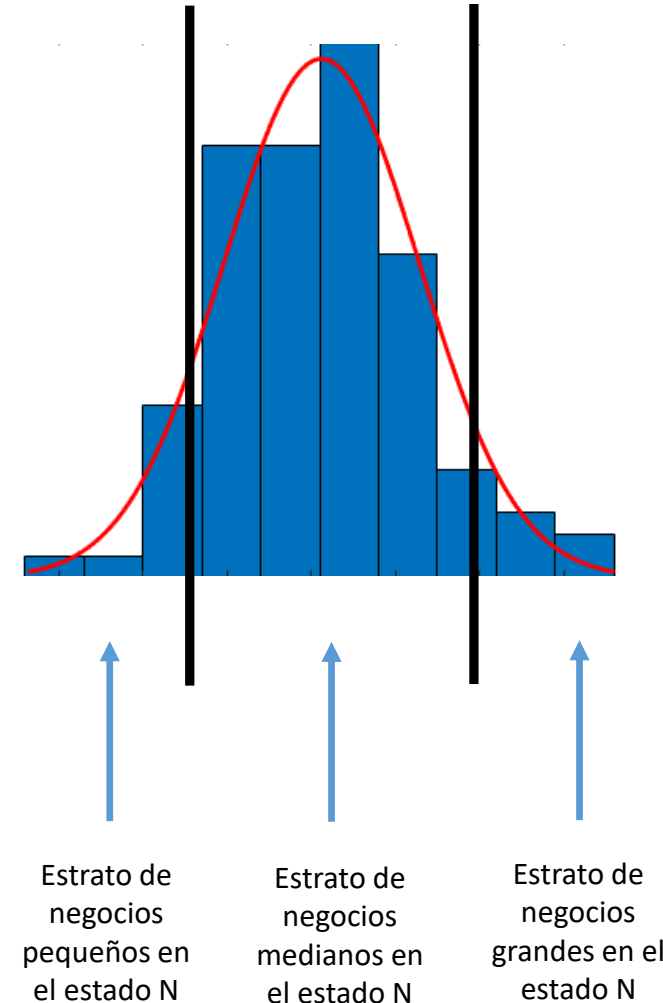
Estado 1



Estado 2

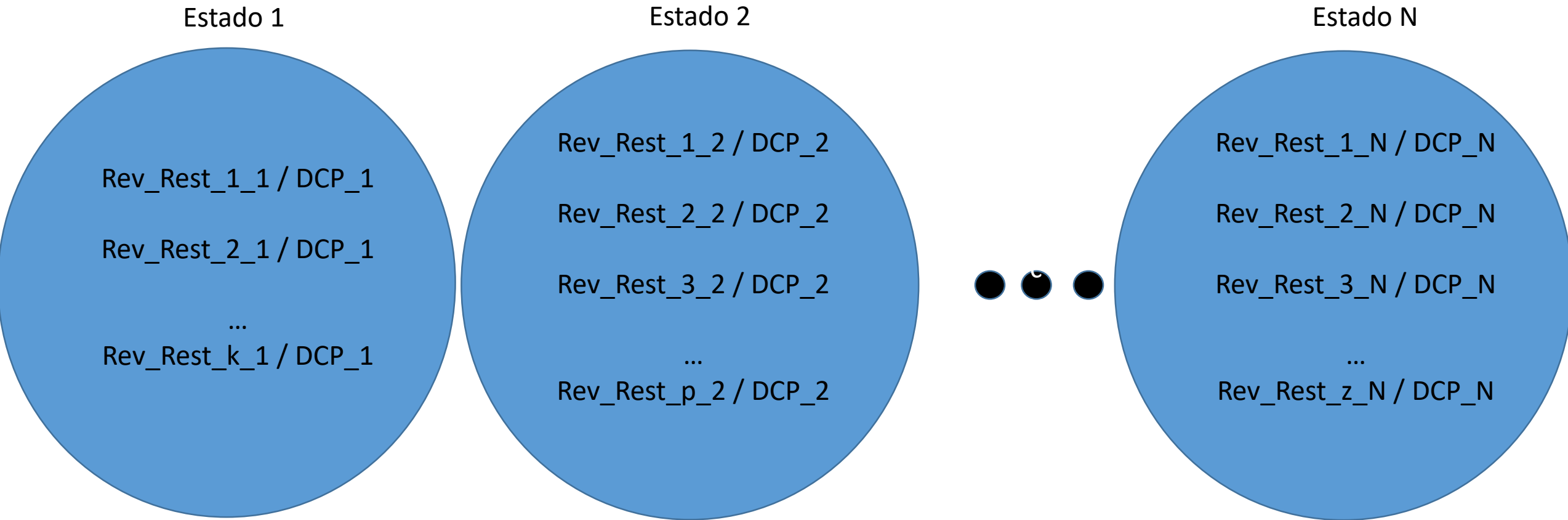


Estado N



Teniendo estos seccionamiento podemos estratificar a los conjuntos de comercio según su “tamaño”, pero este aun cuando es una métrica que nos permite hacer consideraciones dentro del estado. Aun no podemos hacer comparaciones de “consumo” por estrato, entre estados, pues habría que normalizar los datos a efecto de minimizar la particularización de cada uno.

MAXIMO NIVEL DE CONSUMO



Al mismo tiempo, cuando estamos calculando el DCP de cada estado, podemos usar este valor “constante” para evaluar el posicionamiento de cada comercio dentro del estado, generando un índice que se comportara igual que las revisiones en su forma nominal, pero con el valor agregado de poder compara este índice con el índice de otro comercio en otro estado. Es decir, es posible identificar el comercio con mayor consumo en cada uno de los estados y entre ellos, poder identificar el de mayor consumo entre todos. Una elección basada en un criterio relativo dependiente de las condiciones en las que cada comercio opera.

MAXIMO NIVEL DE CONSUMO

Estado

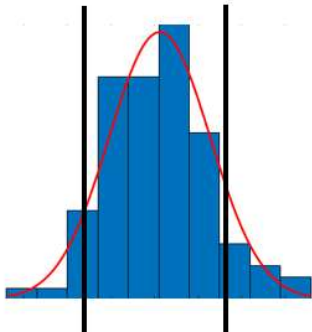
$\text{Rev_Rest_1} / \text{DCP}$

$\text{Rev_Rest_2} / \text{DCP}$

$\text{Rev_Rest_3} / \text{DCP}$

...

$\text{Rev_Rest_p} / \text{DCP}$

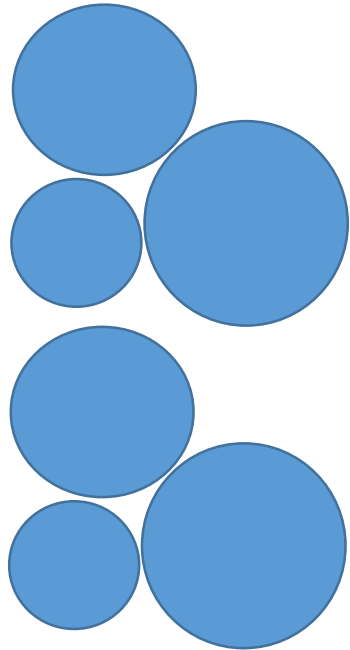


El DCP (Densidad de Consumo de la Población por Restaurante), es un valor único asociado a cada estado.

De esta manera si queremos identificar como es el consumo de un comercio con respecto al estado en el cual se encuentra, será entonces necesario asociar el “count_reviews” de cada comercio con el parámetro de densidad de consumo del estado.

Con este parámetro podemos identificar realmente y de manera proporcional, cual es el comercio con mejor consumo en cada estado. Y en consecuencia poder comparar este parámetro con el resto de los estados. Inclusive hacer comparación de este parámetro entre estados según el estrato de tamaño de negocios.

RESUMEN PARCIAL (índice de concentración, índice de consumo potencial)



Concentración de restaurantes por estado (**CRE**).

Tenemos un KPI que nos permite entender como es la concentración de restaurantes en cada estado. No es un KPI relacionado a los factores internos del cliente, esta mas bien relacionado a factores externos que afectan al cliente. Esta métrica, incide en la decisión de inversión.

$$CRE = \frac{\text{Cant. Restaurantes}}{\text{Tamaño ter. Estado}}$$

Densidad de consumo de la Población por Estado (**DCP**).

Tenemos un KPI que nos permite entender como se relaciona la población con respecto al cumulo de restaurantes en el estado, esto con la finalidad de inferir el potencial de consumo. De manera similar, esta es una métrica de factor externo que impacta en la decisión del cliente.

$$DCP = \frac{\text{Población}}{\text{Cant_rest}}$$

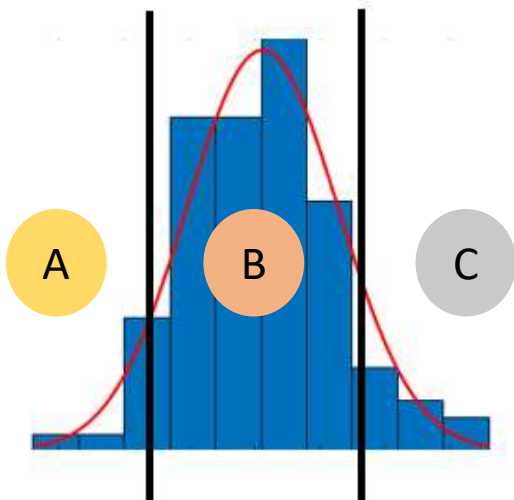
CRE_1
DCP_1

CRE_2
DCP_2

CRE_3
DCP_3

CRE_n
DCP_n

RESUMEN PARCIAL (Índice de consumo estratificado por tamaño del comercio)



C

Comercios sobre los cuales se infiere alta afluencia de clientes (ventas, consumo). 15% en una distribución teórica normal.

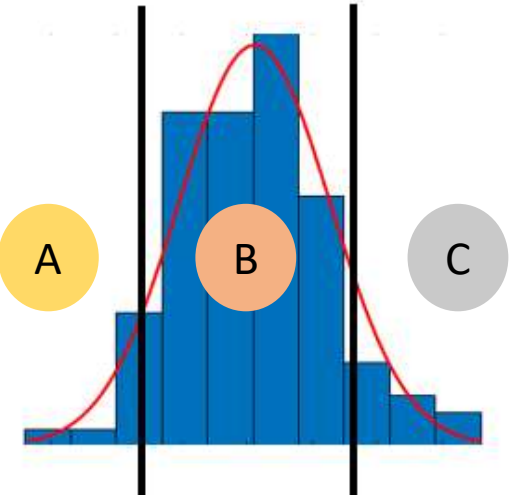
B

Comercios sobre los cuales se infiere una afluencia de rango promedio de clientes (ventas, consumo). 70% en una distribución teórica normal.

A

Comercios sobre los cuales se infiere una baja afluencia de clientes (ventas, consumo). 15% en una distribución teórica normal.

RESUMEN PARCIAL (Índice de consumo estratificado por tamaño del comercio)



- C** Comercios sobre los cuales se infiere alta afluencia de clientes (ventas, consumo). 15% en una distribución teórica normal.
- B** Comercios sobre los cuales se infiere una afluencia de rango promedio de clientes (ventas, consumo). 70% en una distribución teórica normal.
- A** Comercios sobre los cuales se infiere una baja afluencia de clientes (ventas, consumo). 15% en una distribución teórica normal.

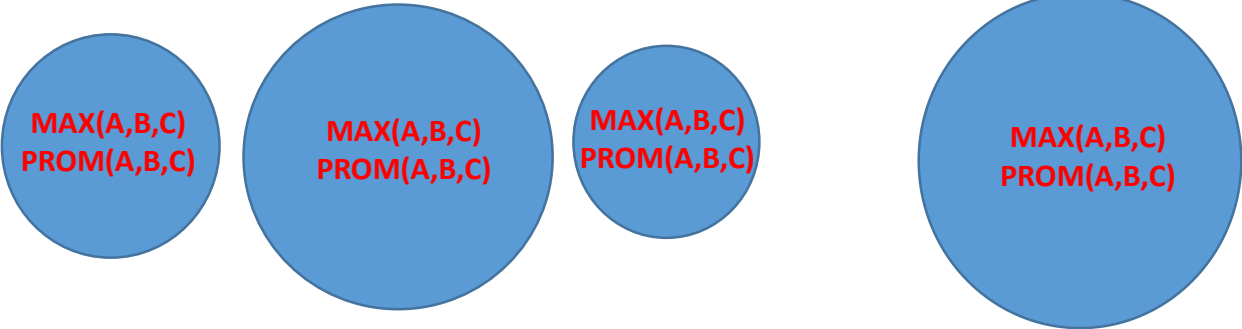
Densidad de Consumo del comercio (Según el estrato de tamaño), por estado (**DCC**).
Tenemos un conjunto de KPIs que proveen información relacionada a como es el nivel de consumo según la estratificación de comercios por “tamaño”. El índice construido permite comparar como es el consumo máximo y promedio en cada estrato de comercios según su tamaño, entre estados.

Sea Dre, la Densidad de revisiones en el restaurante “r” dentro del estado “e”.

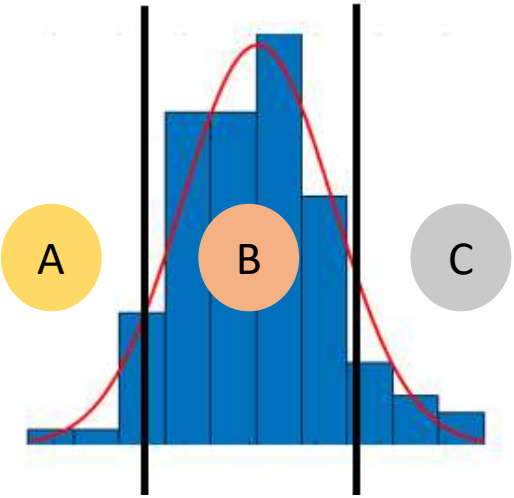
$$Dre = \frac{Rev_Rest_r}{DCP_e}$$

Máximo del índice de consumo por comercio **MAX** { D1a , D2a , D3a , ... , Dna }

Promedio del índice de consumo por comercio **PROM** { D1a , D2a , D3a , ... , Dna }



RESUMEN PARCIAL (Índice de percepción por estrato de tamaño del comercio)

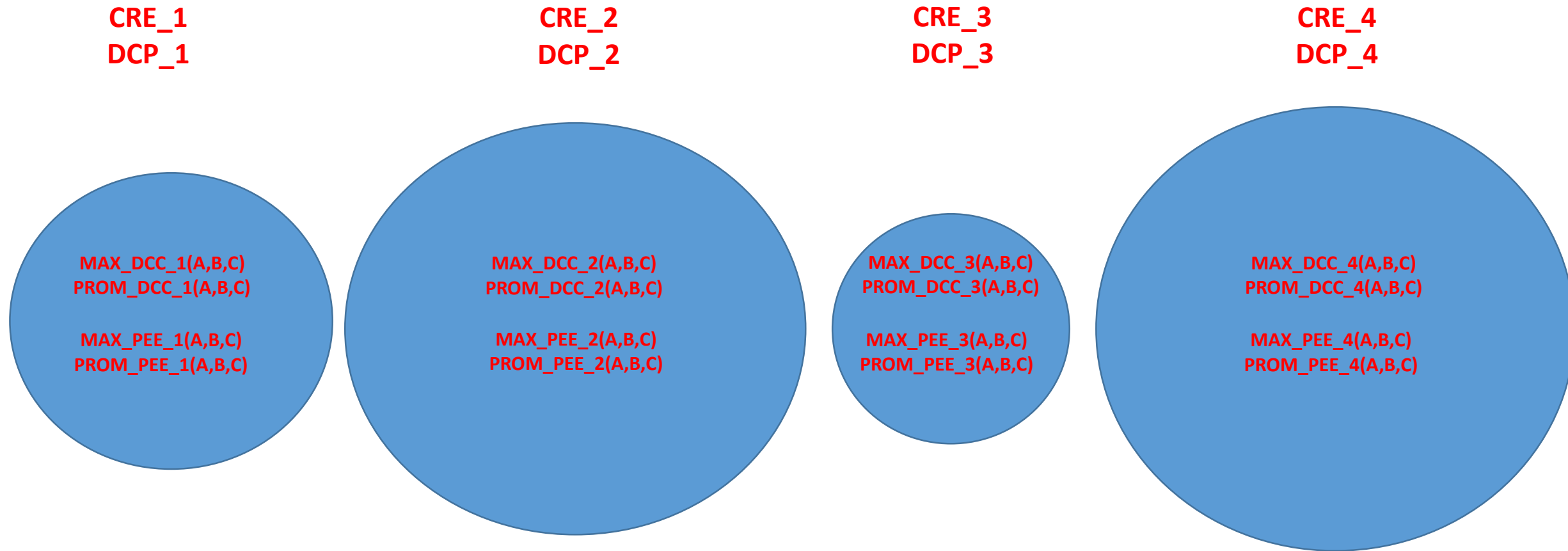


- C** Comercios sobre los cuales se infiere alta afluencia de clientes (ventas, consumo). 15% en una distribución teórica normal.
- B** Comercios sobre los cuales se infiere una afluencia de rango promedio de clientes (ventas, consumo). 70% en una distribución teórica normal.
- A** Comercios sobre los cuales se infiere una baja afluencia de clientes (ventas, consumo). 15% en una distribución teórica normal.

Percepción (Según el estrato de tamaño), por estado (**PEE**).
Tenemos un conjunto de KPIs que proveen información relacionada a como es percibida la calidad de servicio de los comercios según el estrato de tamaño al cual ha sido clasificado.

Estado 1			
MAX(PA1)	{ P1A , P2A , P3A , ... , PnA }	PROM(PA1)	{ P1A , P2A , P3A , ... , PnA }
MAX(PB1)	{ P1B , P2B , P3B , ... , PnB }	PROM(PB1)	{ P1B , P2B , P3B , ... , PnB }
MAX(PC1)	{ P1C , P2C , P3C , ... , PnC }	PROM(PC1)	{ P1C , P2C , P3C , ... , PnC }

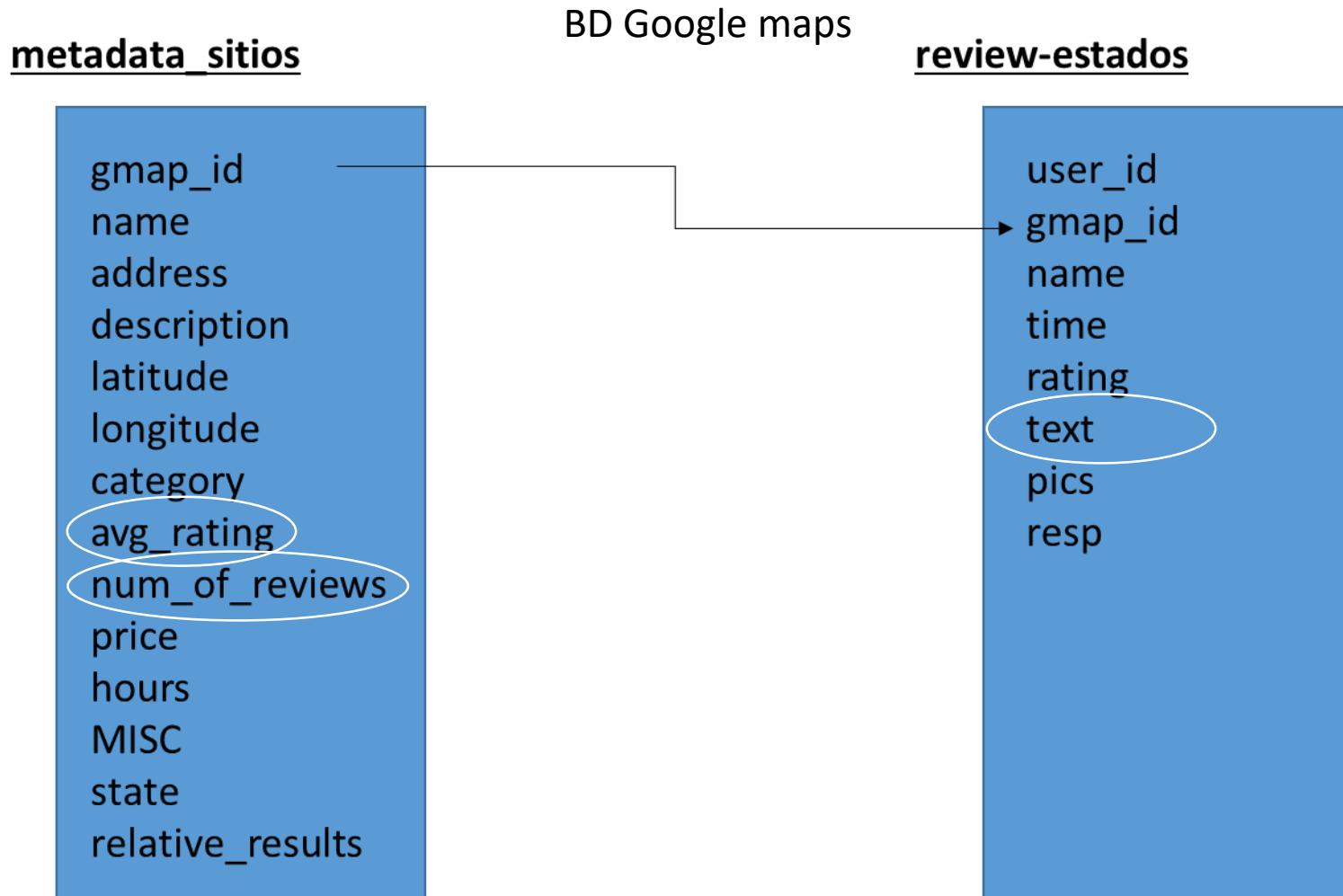
Compendio total de KPIs disponibles)



Los KPIs descriptores de cada estado CRE y DCP dependen de información adicional, Tamaño de territorio y Población, respectivamente.

Los KPIs internos de cada estado están diseñados para poder ser comparados con los mismos parámetros de otros estados.

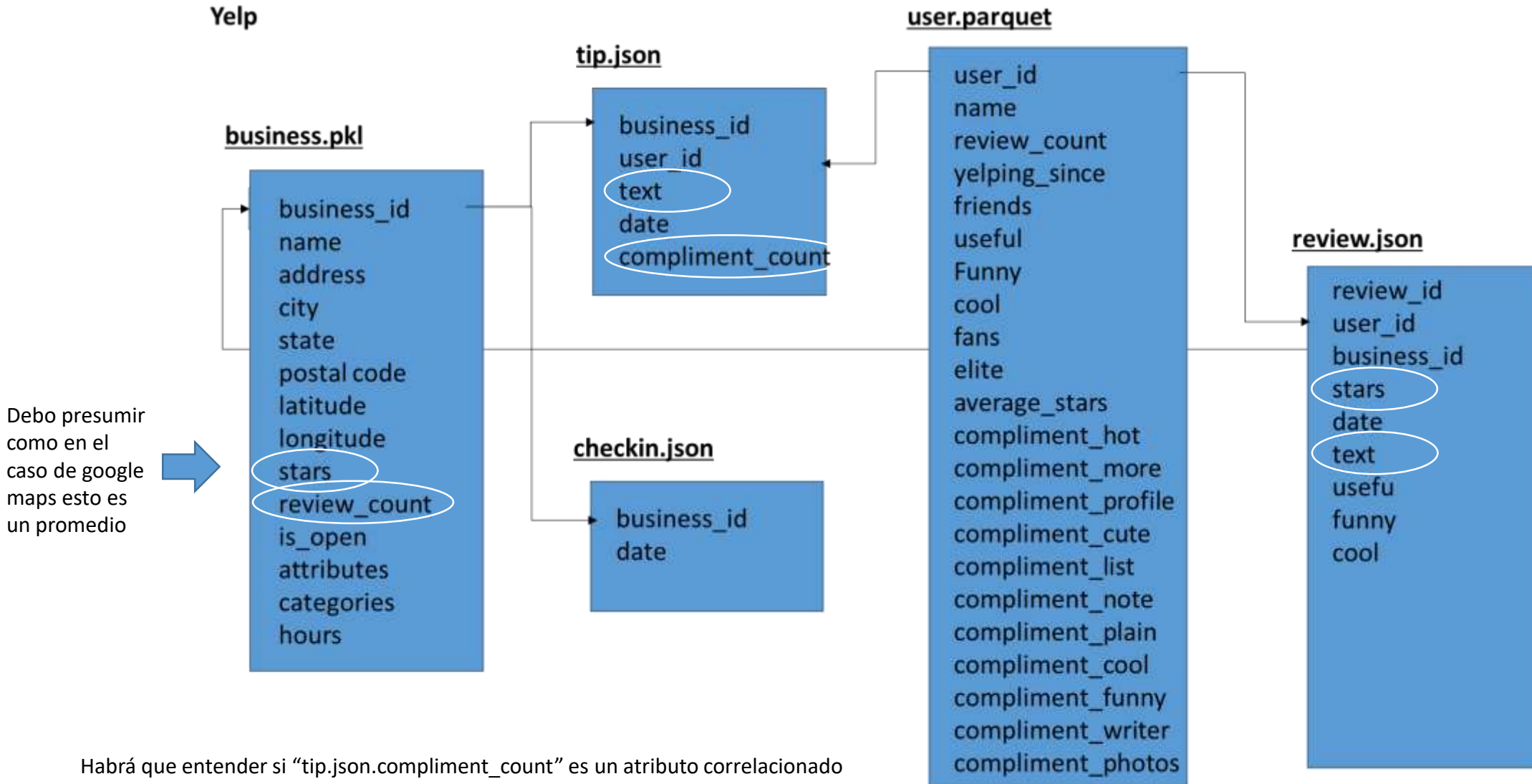
Como obtenemos o generamos las métricas planteadas anteriormente.



A efectos de hacer análisis de consumo, el atributo “num_of_reviews” de la tabla “metadata_sitios” debería ser suficiente, además es necesario que sea dicho atributo pues no hay otro campo útil.

En cuanto al análisis de percepción, esta disponible el atributo “avg_rating” de la misma tabla. En este respecto la siguiente opción es mediante un análisis de palabras en el campo “text” de la tabla review-estados.

Como obtenemos o generamos las métricas planteadas anteriormente.



Habrà que entender si “tip.json.compliment_count” es un atributo correlacionado a “business.pkl.review_count”.

Si resultan positivamente correlacionados, tomaremos el atributo mas completo.

Si los datos están completos y suficientemente limpios en las tablas metadata_sitios de google maps y business.pkl de Yelp, Podremos desarrollar las métricas planteadas y la mayoría del análisis útil.

En dicho caso, el valor que aportan las tablas “reviews” de ambas bases de datos radica en la posibilidad de hacer análisis de tiempo, es decir, análisis de tendencias, por ejemplo, del comportamiento de los usuarios con respecto a los comercios desde el punto de vista de su participación (afluencia) o tendencia en el cambio de la calificación por sector comercial (percepción en el tiempo).