



El futuro de la evaluación programática en la era de reinversión del gobierno

Las instituciones gubernamentales y sus comités de asesores, necesitan información para tomar decisiones acerca de los programas que allí se manejan. Esta información, que se encarga de revelar si un programa está funcionando, generalmente se puede clasificar en tres categorías: descriptiva, acerca de la implementación y acerca de los efectos.

La información descriptiva nos dice en qué consiste un programa y lo que hace. La información acerca de la implementación cubre la parte de los procedimientos, o como el programa se está desarrollando. En el pasado, la información acerca de la implementación hacía referencia acerca del cumplimiento de los requerimientos y procedimientos - más no acerca de los resultados de los programas. Pero hoy en día, el financiamiento de programas está ligado más directamente a los objetivos cuantitativos y a los resultados esperados. Hoy en día es de gran importancia la implementación de planes de evaluación basados en las mediciones para todas las actividades relacionadas con programas del Estado. Siendo más específicos, en la actualidad para asegurar el éxito de un programa es vital:

- Establecer metas de ejecución que definan el programa
- Expresar estas metas en forma objetiva, cuantificable y medible
- Establecer indicadores de gestión que midan o evalúen los niveles de servicio y los resultados
- Ofrecer bases para comparar los resultados actuales del programa con las metas de ejecución establecidas.

Recientemente en Estados Unidos, un reporte proveniente de una oficina de contaduría examinó los estudios de efectividad de diferentes programas de adiestramiento para obtener trabajo. Las técnicas usadas en el reporte incluyen medidas cuantitativas - como las diferencias en términos salariales entre las personas entrenadas, y los que no lo fueron - que pudieran servir como modelo para implementación de programas similares en el futuro.

En este artículo, nos vamos a concentrar en 2 programas, cuya efectividad se mide a través de la información cuantitativa - gran parte de la cual fue recolectada usando SPSS. Trabajar con información cuantitativa ofrece grandes beneficios, sobre todo en esta época donde cada vez se evalúa con mayor detenimiento los programas y su efectividad.

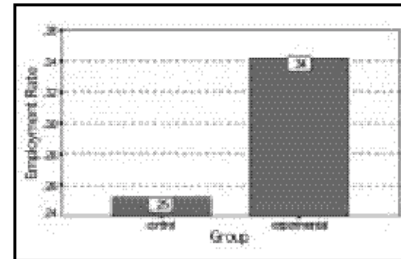
Hoy, el financiamiento de programas está atado directamente a los objetivos cuantitativos de los programas y a los efectos cuantificables

Programa 1

El propósito del primer programa evaluado era ayudar a los solicitantes a encontrar empleo, aumentar las ganancias y reducir los costos de la asistencia social.

Metodología. Un conjunto de personas pertenecientes a un grupo de familias que tienen hijos y reciben ayuda económica fue seleccionado de manera aleatoria y dividido en grupos más pequeños. El primer grupo (el grupo experimental) era obligado a participar en el programa que se está evaluando. El segundo grupo (grupo de control) no tenía obligación de participar en el programa, y tenía libertad para acogerse a otros servicios comunitarios. Al final de un período de dos años, se usaron dos medidas de resultado: tasas de empleo y sueldo promedio. Las diferencias entre los dos grupos con respecto a las medidas de resultado se tomarían como los efectos estimados del programa.

Hallazgos. Los gráficos que resumen los resultados de los análisis estadísticos realizados son la mejor



forma de examinar el comportamiento de los dos grupos en lo referente a la tasa de empleo. La Figura 1 muestra un gráfico de barras de la tasa de empleo.

Figura 1. El gráfico de barras de SPSS muestra la diferencia entre las tasas de empleo para los dos grupos.

Podríamos haber obtenido otra visión de los resultados en términos de empleo a partir de una tabulación

cruzada. Este procedimiento estadístico analiza la asociación entre factores, en este caso la pertenencia a un grupo y el empleo. (Las cifras son hipotéticas, pero representan un resultado posible del programa). La información que aparece en la margen inferior del resultado de la tabulación cruzada (Figura 2) muestra que en general, el 70.7 por ciento de los 474 casos estudiados que reciben ayuda económica están desempleados. Por lo que se esperaría que si se pertenece a determinado grupo (por ejemplo: participación en el programa evaluado) y esto no tiene relación con las tasas de empleo, por lo tanto, 70.7 por ciento de los que pertenecen al grupo de control (o 182.3 personas, la cifra esperada en la celda en la parte superior izquierda) deberían ser los desempleados. Pero este no es el caso.

		Count				
		Unp Val	Emp Val	unemployed employed		
		Row Tot	Row Tot			Row Total
GROUP				0	1	
control	0	182	65			258
		382.3	73.7			25.4%
		74.8%	23.3%			
experimental	1	182	74			256
		352.7	63.3			45.6%
		65.7%	34.3%			
Column		335	139			474
Total		70.7%	29.3%			100.0%

Figura 2. Acá se puede apreciar que el 74.8 % de los pertenecientes al grupo de control estaban desempleados. En contraste, sólo 65.7% de los del grupo experimental estaban sin empleo

Los receptores de ayuda inscritos en el programa ingresaron al mercado de trabajo a una tasa mayor



En realidad, 193 miembros del grupo de control, (los que no participan en el programa que se está evaluando) eran desempleados. Esta misma discrepancia entre los resultados esperados y los resultados obtenidos se presenta en cada una de las cuatro celdas de tablas cruzadas. Esto indica que hay una asociación entre las tasas de empleo y la pertenencia a determinado grupo, y la participación en el programa en cuestión tiene un efecto positivo en el estatus de empleo de los participantes.

Pero la verdadera pregunta es si se pueden esperar los mismos resultados si el programa se aplica a otras personas, por ejemplo la población de Estados Unidos como total.

Chi-Square	Value	DF	Significance
Pearson	4.66200	1	.03084

Figura 3. Esta es una parte del resultado de la tabulación cruzada; indica que hay una relación significativa entre la pertenencia a un grupo y el estatus laboral.

La Figura 3 es parte del resultado obtenido de la tabulación cruzada. El valor de la cantidad (value amount) es una medida del grado en el que cada una de las cifras esperadas de las celdas de las tabulaciones cruzadas difieren de las cifras observadas en las mismas. En otras palabras, este es grado de asociación entre la pertenencia a un grupo y la tasa de empleo. El nivel de significancia de este valor indica si usted puede inferir un grado similar de asociación para una población más grande. Este valor de significancia puede ser interpretado cómo: si la participación en el programa no tuviera efecto alguno en la probabilidad de estar empleado en la población de los Estados Unidos como un todo, entonces la probabilidad arrojada en los estudios realizados sobre la muestra estarían tan sólo un poco por encima de 3 por ciento (.03084). La conclusión? El programa parece estar arrojando los resultados esperados, los receptores de la ayuda ofrecida y que estaban inscritos en el programa en cuestión entraron a formar parte de la fuerza laboral en una tasa más alta que aquellos que no estaban participando en este programa. De hecho, usted podría usar una secuencia de pasos similar para evaluar los efectos del programa en los niveles salariales, tal y como lo hacemos a continuación en el Programa 2.

Programa 2

El objetivo de este segundo programa era similar al primero; ambos estaban diseñados para ayudar económicamente a las

La tabulación cruzada analiza la asociación entre factores



personas en desventaja a competir en la fuerza de trabajo y de esta forma reducir la necesidad de asistencia social.

Metodología. Los aplicantes al programa alrededor del país fueron asignados de manera aleatoria al grupo de tratamiento (con acceso al programa) o al grupo de control (sin acceso al programa) a lo largo de un período de dos años. Los resultados de ganancias fueron recolectados a lo largo de los 18 primeros meses luego de haber determinado cuatro grupos demográficos - hombres y mujeres adultos (22 años o más) y hombres y mujeres jóvenes, graduados del colegio (edades entre los 16 y los 21 años).

Hallazgos. De nuevo, las cifras en el siguiente ejemplo son hipotéticas pero posibles. En esta oportunidad, la evaluación busca estudiar segmentos específicos de la población objetivo para que los resultados del estudio y las consecuentes mejoras necesarias a los programas se enfoquen de mejor manera. Por ejemplo, la Figura 4 muestra la distribución de los niveles de salario mensual para hombres en un gráfico llamado diagrama de cajas. La caja para cada grupo es la mediana, los bigotes y los valores atípicos. La línea más oscura en el gráfico es la mediana actual del salario. Los bigotes por encima y por debajo indican el spread de los salarios por fuera de la mediana.

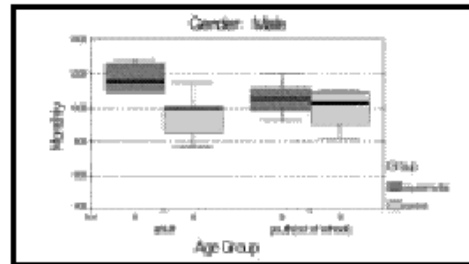


Figura 4. Este diagrama de cajas muestra la distribución de los niveles mensuales de salario para los hombres

El gráfico indica que para los hombres, el impacto del programa en los adultos tiene mayor significancia que en los jóvenes graduados del colegio. Sin embargo, el mismo resultado no aplica para las mujeres. En la Figura 5 se aprecia el efecto para los dos grupos de edades. Hay que notar que la evaluación fue

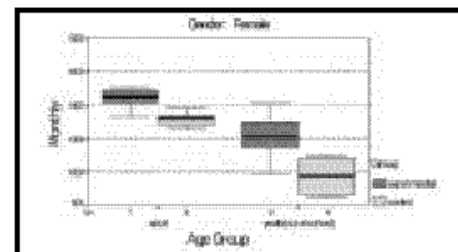


Figura 5. Este diagrama de cajas parece mostrar que hay un impacto para los dos grupos

realizada usando una muestra de la población objetivo. Nosotros construimos gráficos a partir de estas muestras, por lo que la evidencia gráfica sólo se aplica a la muestra. A continuación, necesitamos una manera de inferir resultados similares acerca de la población objetivo a partir de la cual se tomó la muestra. Es acá donde la estadística entra en juego de nuevo. Ahora queremos saber si la diferencia gráfica en el nivel de salario para las mujeres adultas, encontrado en la Figura 5 para un tamaño de muestra de 16 puede ser generalizado de manera razonable para la población.

Para esto, usamos un procedimiento estadístico

Variable	Number of Cases	Mean	SD	SE of Mean
SALARY monthly salary				
exper	7	1458.3216	83.711	31.640
control	9	1313.7035	82.620	27.540
Mean Difference = 144.6181				

Figura 6. Esta prueba t de SPSS muestra si la diferencia gráfica de la Figura 5 puede ser generalizada al total de la población.

Las encuestas de satisfacción de clientes ofrecen otra manera de cuantificar los logros realizados



llamado prueba t (para los resultados, vea la Figura 6). La estadística descriptiva en la parte superior de los resultados de la prueba t, describen la muestra de las 16 mujeres. Los estadísticos también refuerzan los resultados anteriores (los miembros del grupo experimental tienen salarios mensuales más altos). La información en la parte de abajo del gráfico (la prueba t para la Igualdad de las Medias) nos da la información generalizada que queremos. Hay dos líneas desplegadas (igual y desigual) dependiendo de las cuales podemos asumir que la varianza de los salarios en los grupos experimental y de control son iguales. El gráfico de la Figura 5 indica que son iguales, por lo que nos enfocamos en la línea titulada igual.

Examine las columnas tituladas 2-Tail-Sig (o Significancia de 2-Colas) y CI for DIFF (Intervalo de Confidencia por Diferencia). Si dentro del total de la población, asumimos que no hay diferencia dentro de los salarios promedios de los dos grupos, entonces la probabilidad de una diferencia tan o más larga que la encontrada en la muestra que tomamos, es de .004. Esta cifra de la Significancia de 2-Colas nos permite rechazar la suposición de la no-existencia de diferencias en la población. También nos permite concluir que podemos generalizar los resultados obtenidos a partir de muestras.

En la nota relacionada, podemos ver que la diferencia entre los salarios promedios de los dos grupos era de \$144.62 dólares. El CI for Diff (Intervalo de Confidencia por Diferencia) nos dice que si teóricamente fuéramos a tomar varias muestras de tamaño de 16 de la población, 95 por ciento de las diferencias entre los salarios promedio de los miembros de los grupos de control y experimental de esas muestras caerían en ese intervalo. Ese intervalo (\$54.81 a \$234.43) no está cerca de incluir una diferencia de cero (0). Por lo que de nuevo, llegamos a la conclusión de que los resultados obtenidos a partir de la muestra pueden ser generalizados para la población.

No todos los efectos son medidos de acuerdo a los niveles de salario, o a las tasas de desempleo. Las encuestas de satisfacción de clientes son otro medio valioso de medir resultados. Adicionalmente, el análisis cuidadoso de los datos recolectados a partir de las encuestas le señala donde puede estar necesitando recursos adicionales para alcanzar los objetivos de su programa.

Resumen

En este artículo se discutió el futuro de la evaluación de programas. Los ejemplos ilustran como las mediciones cuantitativas y los procedimientos estadísticos ayudan a expresar las metas establecidas de desempeño de una manera objetiva, medible y cuantificable. Adicionalmente se cubrió el tema de cómo los indicadores de gestión cuantificables pueden ayudar a evaluar resultados y niveles de servicio de los programas de gobierno.

SPSS Inc. ofrece soluciones que descubren los que sus clientes quieren y predicen lo que van a hacer a continuación. La compañía ofrece soluciones que se encuentran en la intersección del CRM



(Customer Relationship Management) y la inteligencia de negocios, las cuales les permiten a los clientes interactuar con los clientes de manera más rentable. Las soluciones de SPSS integran y analizan datos de mercadeo, de clientes y datos operacionales en mercados verticales claves alrededor del mundo, esto incluye entre otros: telecomunicaciones, salud, sector financiero, manufactura, retail, investigación de mercados y sector público. Con base en Chicago, SPSS tiene más de 40 oficinas alrededor del mundo.

Contáctenos

Para ponerse en contacto con SPSS:

- Visítenos en Internet: www.spss.com/la

SPSS es una marca registrada y otros productos de SPSS nombrados son registrados por SPSS Inc. Los nombres son marcas registradas de sus respectivos dueños