



## **Enseñanza de la Estadística en contexto: Análisis e interpretación de datos oficiales**

José Paúl **Carrasco** Escobar  
Instituto Politécnico Nacional /CICATA-Legaria  
México

[jcarrascoe1200@alumno.ipn.mx](mailto:jcarrascoe1200@alumno.ipn.mx)

José Luis Ángel **Rodríguez** Silva

Instituto Tecnológico y de Estudios Superiores de Monterrey

México

[luis.rodriguez@inegi.gob.mx](mailto:luis.rodriguez@inegi.gob.mx)

### **Resumen**

El aprendizaje moderno de la Estadística está cada vez más ligado al contexto de los datos y a la atención de los requerimientos de los tomadores de decisiones basadas en información confiable y de calidad. La disposición de estadísticas oficiales sobre los temas sociales, económicos y de los registros de las funciones administrativas y de control de los gobiernos es cada vez más común y su aprovechamiento es una labor propia tanto de especialistas como de otros profesionales que cuentan con formación estadística. En esta comunicación se presentan, como parte de un estudio de caso sobre aprendizaje y evaluación del uso de técnicas estadísticas en el contexto del lugar de trabajo, los resultados de la aplicación de ciertas técnicas estadísticas multivariadas a datos oficiales sobre el medio ambiente, considerando el uso del conocimiento estadístico y el desempeño de un egresado de un posgraduado en Estadística Oficial, quien realizó dicho estudio. Se muestra que el contexto es significativo para el aprendizaje y aplicación de la Estadística.

*Palabras clave:* Educación Estadística, Enseñanza de la Estadística, Didáctica de la Estadística, Análisis multivariado, lugar de trabajo, contexto.

### **Antecedentes**

Durante el proceso de enseñanza-aprendizaje, el desarrollo y aplicación de diversas técnicas estadísticas sobre casos reales son excelentes motivadores para que los estudiantes

perciban el valor de aumentar sus conocimientos estadísticos y así mejorar la calidad futura del análisis y desarrollo de políticas públicas y la toma de decisiones que llevan a éstas (véase, por ejemplo, Wild y Pfannkuch, 1999; Moore, 1997; Biehler, 1994). Los datos, cuando éstos son de calidad adecuada, son esenciales tanto en la toma de decisiones del gobierno como para una correcta participación en la educación de la sociedad y en la obtención de logros académicos y profesionales. Ejemplos de estos son las estadísticas de salud (tasas de morbilidad y mortalidad), los índices de criminalidad y encarcelamiento y la información fiscal. Algunos de estos datos provienen de los registros administrativos en poder de autoridades gubernamentales y otros se derivan de la aplicación de diversos instrumentos estadísticos como encuestas y censos, los cuales a menudo están a cargo de la oficina nacional de estadísticas del país bajo consideración o de investigaciones independientes o académicas. En tal línea de pensamiento, resulta esencial que los educadores estadísticos motiven y permitan que los estudiantes tengan las habilidades para evaluar la calidad relativa de los datos oficiales que se usen, así como incentivar el uso de tales estadísticas (Forbes y Janssen, 2014). Los estudiantes deben ser alentados a utilizar, así como a entender las estadísticas oficiales y, por tanto, a evitar su mal uso (Janssen y Sharlee, 2014).

Por otro lado, es necesario que cada país promueva el fomento a su capacidad estadística, entendiéndose esta como *las actividades que fortalezcan las diferentes actividades consideradas como pilares, tales como la promoción para el fortalecimiento de las estadísticas oficiales y el conocimiento, la transferencia y la formación técnica en la recopilación de datos, su análisis y la difusión de estos* (Hermann, 2014).

Para incentivar tal capacidad estadística, diversos autores, como Westbrooke y Ellis (2014) proponen que los conocimientos estadísticos clave se promocionen y se usen en los lugares de trabajo, incluyendo: el entrenamiento en el uso de gráficos para exploración y visualización de datos; el modelado, especialmente los modelos lineales y sus extensiones, con énfasis en el efecto de tamaños de las muestras en lugar de las pruebas; el diseño y análisis adecuado de los datos *pequeños y de los grandes*; y la aplicación de software de estadística de apoyo al entrenamiento. Una de las estrategias para hacer frente a estas necesidades y desafíos en el lugar de trabajo es aumentar las habilidades de los no expertos en estadística en su lugar de trabajo. Sus experiencias refuerzan el énfasis en la literatura sobre la importancia de aplicar diversas técnicas estadísticas con el uso de los datos y problemas reales, con una aplicación directa a los proyectos de trabajo. Finalmente, concluyen que las *grandes ganancias* vienen rápidamente cuando el personal descubre que pueden aplicar las nuevas habilidades y herramientas para llevar a cabo su trabajo más fácil y rápidamente y, con frecuencia, con una calidad superior. Esto significa que las necesidades de formación para el lugar de trabajo integran a los participantes que realizan prácticas y tareas, aumentando sus habilidades y los anima a compartir y demostrar las aplicaciones que los retos que sus actividades laborales les plantean.

Respecto a los laboratorios de computación, existen posturas sobre qué tan apropiado es la enseñanza de aplicaciones estadísticas con GUI (interfaz gráfica de usuario) o, en su lugar, hacer uso de software para cómputo estadístico a través de algoritmos de programación o de líneas de comandos o scripts, como lo que se realiza con ayuda de lenguajes como *R* o *Scilab* ©. Por ejemplo, Navarro (2014) menciona que no es raro encontrar biólogos con un estado de ánimo vacilante, sintiéndose satisfechos con la solución de un problema, porque fue logrado con la ayuda de una interfaz gráfica de usuario inteligente, mientras que la situación real es que fue el resultado de un procedimiento automatizado o encapsulado de *caja negra*. Agrega que este es el

precio que se paga de ser sólo un simple usuario y que los riesgos de la generación de resultados a través de las *cajas negras* se pueden minimizar si el usuario está familiarizado con los principios conceptuales y procesales de la programación y la informática, pertenecientes a lo que llama *la computación del siglo XXI*. Concluye que es urgente que los científicos, y en términos generales, los usuarios, sean reeducados en los temas informáticos de la ciencia en general, y de la estadística, en particular.

Para contribuir a la formación en el lugar de trabajo, Bavdaž, Ograjenšek y Perviz (2014) mencionan que se necesitan datos de las estadísticas oficiales en el lugar de trabajo (para la evaluación comparativa, análisis de mercado, etc.), pero estos han sido, en gran medida, subutilizados. De un estudio estadístico realizado por dichos investigadores se tiene que, en general, se utilizan las estadísticas oficiales con el propósito de familiarizar a los estudiantes con el estado actual de un fenómeno determinado (83%) o para ilustrar su evolución en el tiempo (82%). En menor parte se utilizaron las estadísticas oficiales para explicar la metodología detrás de un indicador elegido (34%) o para explicar técnicamente un método analítico (29%); si lo hicieran, se utilizaría alguno de ellos, principalmente para hacer una combinación con la presentación del estado de un fenómeno y de su evolución. Solamente un 12% utiliza los cuatro enfoques mencionados. Mencionan también que las ideas sobre el uso de las estadísticas oficiales en los programas de estudio de negocios, por ejemplo, al parecer están casi siempre vinculadas a la interpretación y, muy a menudo, a la comprensión conceptual. También se asocian con otros elementos pero es probable que algunos de estos elementos estén estrechamente conectados con el tema del curso (por ejemplo, con la computación estadística). Sobre la base de sus resultados, es posible concluir que sólo un pequeño grupo de los educadores pueden ser caracterizados como usuarios asiduos de las estadísticas oficiales; algunos educadores no los usan en absoluto y otros los utilizan ya sea principalmente para la propia preparación y presentación de las actividades a los estudiantes. El uso más generalizado de las estadísticas oficiales para conferencias es también más pasivo, desde el punto de vista de los estudiantes.

Las actividades en el uso y aplicación de técnicas estadísticas sobre datos oficiales ofrecen entonces a los estudiantes experiencias más directas y con mayor sentido a ellos que los ejemplos que no utilizan tales estadísticas oficiales, lo anterior en términos generales. En tal sentido, pudiera parecer prudente que la mayoría de los educadores que utilizan las estadísticas oficiales ofrecieran a sus estudiantes, a su vez, una mínima oportunidad de utilizarlas en sus propios procesos de toma de decisiones (por ejemplo, en la aplicación de proyectos de inversión, de emprendimiento, para llevar a cabo un análisis fundamental sobre algún tópico que sea de su interés, etc.). Esto trae consigo, de manera dual, la correspondencia a invitar a los proveedores de estadística oficiales a brindar sus experiencias a los estudiantes para que éstos puedan conocer de manera directa la forma en que se desarrollan los procesos de las estadísticas oficiales (diseño conceptual, definición de objetivos y metas, alcance del proyecto, representatividad geográfica, desarrollo del instrumento de captación, captura, validación, tratamiento, explotación y difusión de las estadísticas oficiales), mediante la exposición de ejercicios interactivos y de interés a los propios alumnos. En tal sentido, en México, el Instituto Nacional de Estadística y Geografía tiene ya, en estos momentos, un área que brinda este tipo de asesorías mediante la adherencia, por ejemplo, a la llamada “Cátedra INEGI” y cuya disponibilidad hace más fácil enlazar tanto a los propios estudiantes como a los mismos profesores en la pertinencia y el valor de las estadísticas oficiales para el proceso de enseñanza y aprendizaje, en el marco de sus intereses particulares.

El objetivo del presente artículo, bajo la línea de pensamiento planteada, es el presentar un estudio de caso sobre la forma en que un egresado de la Maestría en Ciencias en Estadística Oficial, trabajador del INEGI, realiza análisis estadístico a datos oficiales, en el contexto de su lugar de trabajo, para generar información que nutra la toma de decisiones en políticas públicas relacionadas con la gestión ambiental.

### **Fundamentación teórica**

Enfocándonos a lo referente a un proceso de formación profesional o especializada, los estudios de casos son utilizados para establecer el aprendizaje de la Estadística en el contexto del mundo real de los estudiantes (Forbes y Keegan, 2014). En los últimos años esto se ha basado en la teoría de pensamiento estadístico, centrada en las estadísticas oficiales, así como en los métodos generales de estadística. Vaughan (2008) afirma que “el aprendizaje en el lugar de trabajo no es sólo un proceso unidireccional, es una interacción entre el lugar de trabajo, el aprendizaje, y el alumno”. Sin embargo, lo que aún no es del todo claro es si dicho proceso didáctico logra, en efecto, un impacto positivo y duradero en el rendimiento de las organizaciones en cuanto a sus procesos institucionales de toma de decisiones o incluso en la mejora de la calidad de la producción estadística, aunque existen en proceso diversas investigaciones para caracterizar este complejo fenómeno (e. g., Carrasco, 2014).

Se ha determinado en otros estudios que el razonamiento estadístico será una de las facetas más importantes de la vida profesional de los egresados dado el mayor nivel de demanda de la utilización de diversas herramientas estadísticas, por no decir simplemente del acentuado requerimiento de mayores niveles de alfabetización estadística básica, pertenecientes a los así llamados pensamiento crítico y el razonamiento estadístico, que aparecen en políticas, prácticas, atributos y habilidades que un estudiante debería obtener en su desarrollo de actividades académicas. Por un lado, la lista de habilidades específicas que debe mostrar el estudiante incluye los siguientes: demostrar conocimiento y la capacidad de aplicar los principios y conceptos científicos; resolver problemas mediante la aplicación de conocimientos y métodos científicos; demostrar capacidad de análisis; demostrar el desarrollo de habilidades para el aprendizaje permanente; entender, evaluar, acceder y revisar críticamente la nueva información, demostrar la capacidad de pensar de forma independiente; demostrar un profundo conocimiento y habilidad en su sujeto de estudio con especialización; acceder efectivamente y utilizar la información relevante para un tema; demostrar habilidades numéricas, trabajar en colaboración en las tareas; y comunicar las ideas y hallazgos de la aplicación de técnicas estadísticas de manera efectiva.

Para los estudiantes de posgrado de estadística, lo anterior se traduce en centrar la atención en el desarrollo de su propia cultura estadística mediante la ejecución de seminarios estadísticos sobre los principales temas e incrementado sus habilidades prácticas de laboratorios de cómputo estadístico y uso de aplicaciones especializadas en estadística, mediante la utilización de diversas plataformas como R, Minitab®, SPSS ©, entre otras. Estos laboratorios proporcionan a los participantes la experiencia inmediata que ayuda a consolidar su aprendizaje en los seminarios. Algunos de esos temas principales han sido: comprensión e interpretación resultados estadísticos; diseño experimental; ANOVA; modelos lineales generalizados; y modelos de efectos mixtos (Brown, David, Moltchanova, Seddon y Harlow, 2014), mismos que figuran ya como asignaturas en diversos programas de posgrado de México y otras latitudes. Aquí nuevamente el contexto y la aplicación de las técnicas estadísticas sobre datos reales es extraordinariamente relevante y ha sido reconocido por autores como Cobb (2007) quien declaró que “el contexto es un elemento

esencial, parte del pensamiento estadístico, y algunos de las peores enseñanzas de la estadística se produce cuando el profesor o libro de texto intenta tratar al contexto como irrelevante”. Así, el contexto permite motivar el aprendizaje en estadística y es importante seguir actualizando estudios de caso, en colaboración con colegas de una amplia gama de disciplinas que aplican la propia Estadística. Por tal motivo, es necesario asegurar que los estudiantes reciban, desde el principio de su formación e incluso después de ésta, las oportunidades de participar en actividades y la toma de decisiones que se espera de ellos al concluir su programa de estudio, situación que sin embargo, y de manera ciertamente lamentable, rara vez se practica en la educación superior. En lugar de ello, predominan las enseñanzas del curso dirigidas por el instructor, usando un método específico (cortando con ello el fomento de la creatividad de los estudiantes) y mediante el uso de algunos ejemplos artificiales de final de libro (Johnson y Shumway, 2014).

Por lo anterior, resulta de gran interés, desde la óptica de la Educación Estadística, el conocer la forma en que se desenvuelven los profesionales de la estadística sobre ambientes contextuales reales en su trabajo, motivo que será de estudio en el presente documento.

### **Diseño y metodología**

En este artículo se comparte el desarrollo de un caso basado en el entorno, que permite a los profesores desafiar a los estudiantes a demostrar su capacidad de aprendizaje al participar con éxito en el análisis de composición abierta, la evaluación y la aplicación de los contenidos específicos de cada conocimiento dentro del mundo real en contextos de consultoría estadística, y lo anterior con la finalidad de proveer a los estudiantes acceso a indicadores de su eficiencia al desarrollarse como nuevos profesionales a través de su programa de estudios, lo cual maximiza las posibilidades de alcanzar niveles profesionales de orden superior.

Por ejemplo, en la taxonomía de Bloom, los objetivos de aprendizaje de conocimientos conceptuales y procedimentales son más fácilmente identificados dentro de cursos de preparación individuales, donde se les pide a los estudiantes a recordar, entender y aplicar. Muy a menudo es sólo al final de un programa de estudio de los cursos donde se encuentran el análisis, la evaluación y la síntesis que se requieren. En tal sentido, Razzouk y Razzouk (2008) informan que, mientras el análisis y la síntesis son el núcleo de lo que los estudiantes deben ser capaces de hacer como un graduado, estas son “probablemente algunas de las actividades de aprendizaje menos comprendidas y más tergiversadas” (p. 49). Por lo general, este tipo de objetivos de aprendizaje de orden superior se infunden en las experiencias de culminación, ya que son difíciles de abordar intencionalmente en la mayoría de los cursos de preparación.

La enseñanza basada en casos es un método popular que coloca a los estudiantes frente a situaciones que requieren del pensamiento analítico y promueven la síntesis dentro de un campo profesional de estudio en particular. La instrucción basada en casos, sobre la descripción de los complejos acontecimientos que ocurren típicamente en el mundo real da a los estudiantes contextos donde revisan, en el aumento de la complejidad, escenarios posibles. Sin embargo, mientras que la instrucción basada en casos es un método de enseñanza que se encuentra en muchas profesiones y clases de culminación académica, no es comúnmente usada como una herramienta integrada para apoyar el aprendizaje del estudiante desde el inicio del plan de estudios académicos.

El método general para el estudio que reporta este artículo, inicia con la propuesta a un recién egresado de un programa de posgrado en Estadística oficial, al que se le denomina como

el *especialista*, de seleccionar un tema sobre el cual existan publicadas estadísticas oficiales, las cuales sean susceptibles de un análisis estadístico que permita obtener indicadores o conclusiones que sean la base de los elementos de toma de decisiones de políticas públicas en el ámbito del tema seleccionado. En tal sentido, dicho especialista expuso el contexto de los datos y del análisis, así como de la pertinencia del estudio y de las técnicas estadísticas usadas a partir de su acervo de conocimientos individuales y de su capacidad para utilizarlas eficientemente. Los hallazgos se presentan en las siguientes secciones.

### Resultados

A continuación se hace una síntesis del reporte obtenido del especialista recién formado.

#### Contexto sobre los antecedentes de los residuos sólidos urbanos

El especialista comentó acerca del contexto de su problema estadístico lo siguiente:

“La gestión de los *Residuos Sólidos Urbanos*, RSU (o basura doméstica, aunque sin la precisión que tiene el primer concepto), es uno de los fenómenos que más impacto tienen sobre el medio ambiente y sus recursos. Lo anterior es cierto puesto que, por un lado, debido a la creciente demanda que tienen los seres humanos sobre sus satisfactores (que muchos de ellos al final se convierten precisamente en RSU) se tiene un mayor consumo de los recursos naturales, teniendo en consecuencia una afectación a los recursos naturales; por otro lado, al hacer la disposición final de los RSU, y debido tanto a su gran volumen (como será comentado a mayor detalle posteriormente) así como al hecho de que en diversas ocasiones no se cumplen con los estándares y normas al respecto, se hace una afectación al medio ambiente. Esto trae, al final, graves problemas en materia de salud pública, utilización de espacios y recursos y contaminación ambiental, entre otros. Son numerosos los ejemplos en los cuales, por una falta de cuidado en alguno de los múltiples procesos del sistema de los RSU, se tienen consecuencias de impacto al medio ambiente, algunos de los cuales se comentan.

En sus aspectos más simples, la gestión de los RSU implica tres grandes etapas: generación; recolección; y disposición final. Sin embargo, para dimensionar lo complejo de la gestión de los RSU, en la *Figura 1* se presenta el diagrama de flujo de un sistema de manejo de residuos sólidos diferenciado, en el cual resulta claro que se encuentran involucrados una gran cantidad de factores en dicho proceso, por lo que resulta imprescindible el fomentar un uso adecuado en cada una de tales instancias para lograr una mayor gestión operativa al respecto.”

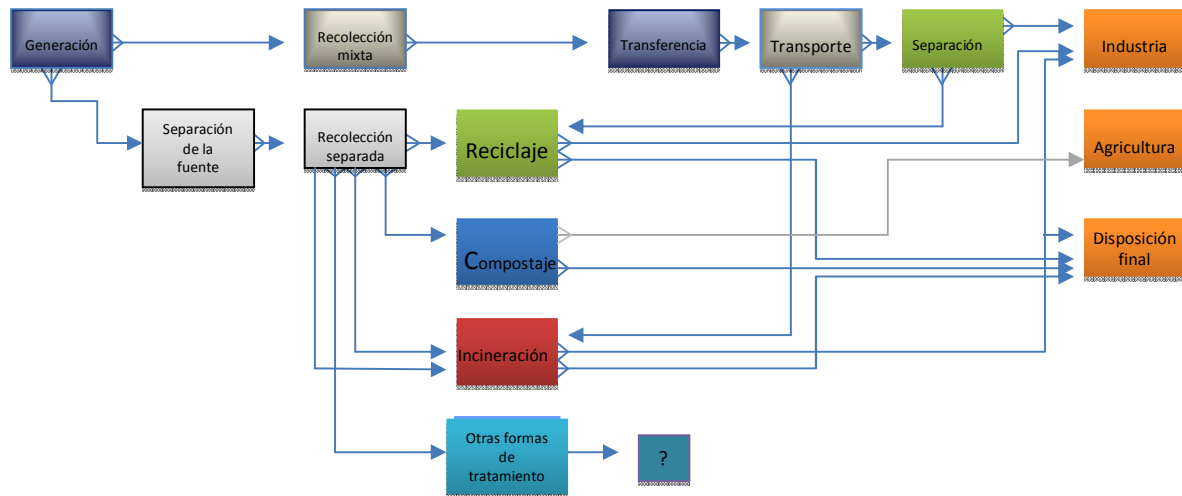


Figura 1. Diagrama de flujo de un sistema de manejo de residuos sólidos diferenciado.

Fuente: Diseño del especialista inspirado en SEMARNAT (2007).

### Técnicas estadísticas multivariadas

Las dos técnicas estadísticas multivariadas que utilizó el especialista fueron el Análisis de Componentes Principales (ACP) y el Análisis de Conglomerados (AC) que, para sus fines, se basó en las referencias siguientes: para el caso del ACP tomó a Johnson, 2000, pp. 93-146; Johnson y Wichern, 2002, pp. 426-476; Manly, 1986, pp. 59-71; para el caso del AC tomó a Johnson 2000, pp. 319-396; Johnson y Wichern 2002, pp. 668-745 y Manly 1986, pp. 100-113. Para el análisis numérico de datos se basó tanto en el uso de Minitab® como de R (2014).

### Aplicación de las ACP y Análisis de Correlaciones a un subconjunto de datos de RSU

El especialista comentó acerca de su justificación para la utilización de los datos analizados lo siguiente:

“El Instituto Nacional de Estadística y Geografía (INEGI) de México ha realizado un gran esfuerzo por acopiar información de RSU mediante la adherencia de un módulo en la materia en los así llamados Censos Nacionales de Gobierno, y mediante el cual se ha logrado generar datos de gran relevancia, por los siguientes motivos: antes de este ejercicio de Estadística Oficial no se tenían, de manera abundante y profusa, datos al nivel de detalle con los que ahora se cuentan, y varios de ellos provenían de estimaciones -véase, por ejemplo, SEMARNAT (2012). Así, el INEGI al momento cuenta con, y pone a disposición de los usuarios en general, un vasto conjunto de datos acerca del fenómeno de los RSU. Para el ejercicio mostrado en este trabajo, se acopiaron diversas variables asociadas de los RSU, según está dispuesta la información en el sitio web del INEGI, para determinar niveles de asociación multivariados en las distintas fases de la gestión integral de residuos. A su vez, no se consideraron variables que aunque son de interés, no contaban con registros (por ejemplo, algunas variables de selección de RSU, como PET u otras). Así, en total, se consideraron 20 variables de la siguiente manera:

- VAR<sub>1</sub> = Proporción de municipios y delegaciones con centros de acopio respecto al total de municipios en la Entidad Federativa.
- VAR<sub>2</sub> = Proporción del total municipios y delegaciones con centros de acopio respecto al total nacional de centros de acopio.

- VAR<sub>3</sub> = Proporción de municipios y delegaciones con estaciones de transferencia, sólo almacenamiento temporal y traspaso respecto al total de municipios en la Entidad Federativa.
- VAR<sub>4</sub> = Proporción de municipios y delegaciones con estaciones de transferencia, compactación y/o selección de materiales respecto al total de municipios en la Entidad Federativa.
- VAR<sub>5</sub> = Proporción de municipios y delegaciones con estudios sobre la generación de residuos sólidos urbanos respecto al total de municipios en la Entidad Federativa.
- VAR<sub>6</sub> = Proporción de municipios y delegaciones con programas locales orientados a la gestión integral de los residuos sólidos urbanos respecto al total de municipios en la Entidad Federativa.
- VAR<sub>7</sub> = Proporción de municipios con estaciones de transferencia respecto al total de municipios de la Entidad Federativa.
- VAR<sub>8</sub> = Proporción del total de municipios con estaciones de transferencia respecto al total nacional.
- VAR<sub>9</sub> = Proporción de vehículos utilizados para la recolección de residuos sólidos urbanos con compactador respecto al total en la Entidad Federativa.
- VAR<sub>10</sub> = Proporción de vehículos utilizados para la recolección de residuos sólidos urbanos con caja abierta respecto al total en la Entidad Federativa.
- VAR<sub>11</sub> = Proporción de otros vehículos utilizados para la recolección de residuos sólidos urbanos y otro tipo de vehículos respecto al total en la Entidad Federativa.
- VAR<sub>12</sub> = Proporción del total municipios y delegaciones con disponibilidad de servicios relacionados con los residuos sólidos urbanos respecto al total en la Entidad Federativa.
- VAR<sub>13</sub> = Proporción de municipios y delegaciones con disponibilidad de servicios de sólo recolección y disposición final de residuos sólidos urbanos respecto al total en la Entidad Federativa.
- VAR<sub>14</sub> = Proporción de municipios y delegaciones con disponibilidad de servicios de recolección, disposición final y tratamiento de residuos sólidos urbanos respecto al total en la Entidad Federativa.
- VAR<sub>15</sub> = Proporción del promedio diario de pesaje de residuos sólidos urbanos recolectados respecto al total en la Entidad Federativa.
- VAR<sub>16</sub> = Proporción del promedio diario por vehículo/capacidad/viajes de residuos sólidos urbanos recolectados respecto al total en la Entidad Federativa.
- VAR<sub>17</sub> = Proporción del promedio diario de otros métodos de obtención de residuos sólidos urbanos recolectados respecto al total en la Entidad Federativa.
- VAR<sub>18</sub> = Proporción de municipios y delegaciones con servicio de recolección y disposición final con reglamento relacionado con los residuos sólidos urbanos respecto al total en la Entidad Federativa.
- VAR<sub>19</sub> = Proporción de sitios de disposición final tipo relleno sanitario reportados como destino de los residuos sólidos urbanos respecto al total en la Entidad Federativa.
- VAR<sub>20</sub> = Proporción de sitios de disposición final tipo tiradero a cielo abierto reportados como destino de los residuos sólidos urbanos respecto al total en la Entidad Federativa”.

### **Análisis de Correlaciones**

El primer análisis llevado a cabo por el especialista fue el de correlaciones. Comentó él que dicho análisis es fundamental para poder emplear de manera eficaz el Análisis de Componentes



Principales así como el Análisis de Conglomerados. Textualmente comentó lo siguiente: “al llevar a cabo un análisis de correlaciones, se observó que del total de  $(19 \times 20) / 2 = 190$  correlaciones de variables en sus cruces (eliminando las correlaciones triviales), y al contar el número de correlaciones que, en valor absoluto, fueron mayores a 0.7, se encontró que estas últimas fueron 35, por lo que hay una proporción no trivial, casi un 18.42%, que gozan de dicha propiedad. Si el valor de umbral de 0.7 disminuye 0.5, el cual para cierto tipo de análisis se considera un valor alto, la proporción se incrementa a casi un 25%. Más aún, existen correlaciones muy altas, por ejemplo, entre  $VAR_{11}$  y  $VAR_{16}$ , de un orden de 0.99. Estas dos consideraciones justifican, en parte, el que sea razonable llevar a cabo tanto el ACP como el AC”.

### Aplicación del ACP

En la Tabla 1 el especialista presentó los valores de las varianzas explicadas por cada componente principal (CP, subindexada según su orden). Observó que, utilizando la regla de *más un 70% de varianza explicada*, serían necesarias tres componentes principales para explicar nuestro fenómeno de 20 dimensiones. Sin embargo, señaló que si se exigiera poco menos, y empleamos la regla de *un mínimo de 60%* de varianza explicada, bastaría con dos componentes principales. Este punto fue esencial dentro de nuestro estudio, puesto que nos percatamos que este especialista no considera como “inviolables” las reglas dadas en los textos usuales de estadística, sino que las va adecuando según cada caso particular, con la debida medida y explicitación de la variación correspondiente.

El especialista comentó que: “así entonces, parece claro que el “verdadero” problema de los RSU tiene dos a tres dimensiones, y no tanto como 20, que es el número total de variables de proporción que están bajo análisis. La anterior afirmación se refuerza al considerar el diagrama de sedimentación, mostrado en la *Figura 2*, en la cual es claro que *el codo* de la misma se presenta en la segunda componente principal, motivo por el cual nos inclinamos a la consideración de que la dimensionalidad más importante es de 2, en el sentido de que el 61.7% de la varianza acumulada puede explicarse con tal número de componentes principales.”

Tabla 1

*Valores propios, proporción y proporción acumulada de las CP's sobre datos de RSU*

	$CP_1$	$CP_2$	$CP_3$	$CP_4$	$CP_5$	$CP_6$	$CP_7$	$CP_8$
Valor propio	9.5721	2.7735	2.2142	1.5967	1.1128	0.9018	0.4585	0.3828
Proporción	0.479	0.139	0.111	0.08	0.056	0.045	0.023	0.019
Acumulada	0.479	0.617	0.728	0.808	0.863	0.909	0.931	0.951

*Fuente:* Elaboración del especialista con resultados generados en Minitab®

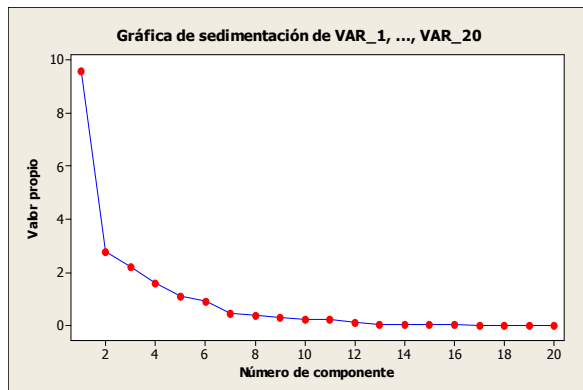


Figura 2. Gráfica de sedimentación del ACP sobre datos de RSU.

Fuente: Elaboración del especialista con resultados generados en Minitab®.

El especialista hizo explícitas las dos primeras CP's de la siguiente forma:

$$Z_1 = 0.294VAR_1 + 0.303VAR_2 + 0.205VAR_3 + 0.024VAR_4 + 0.229VAR_5 + 0.110VAR_6 + 0.265VAR_7 + 0.211VAR_8 + 0.282VAR_9 + 0.165VAR_{10} + 0.295VAR_{11} + 0.008VAR_{12} - 0.301VAR_{13} + 0.302VAR_{14} + 0.082VAR_{15} + 0.295VAR_{16} + 0.302VAR_{17} + 0.096VAR_{18} - 0.005VAR_{19} - 0.174VAR_{20}$$

$$Z_2 = -0.047VAR_1 - 0.145VAR_2 + 0.328VAR_3 + 0.227VAR_4 + 0.119VAR_5 - 0.010VAR_6 + 0.261VAR_7 - 0.014VAR_8 - 0.110VAR_9 - 0.322VAR_{10} - 0.075VAR_{11} + 0.323VAR_{12} - 0.111VAR_{13} - 0.060VAR_{14} + 0.020VAR_{15} - 0.184VAR_{16} + 0.028VAR_{17} + 0.391VAR_{18} + 0.464VAR_{19} - 0.301VAR_{20}$$

Él observó que la primera componente principal adquiere sus mayores valores en las variables 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 13, 14, 16 y 20. Al analizar las primeras, notó que estas se encuentran involucradas con la infraestructura (vehículos y medios para pesar los RSU) en la gestión integral de los RSU. A su vez, observó que los signos negativos se presentan en las variables 13 y 20 se asocian básicamente a la disposición final de los residuos, aunque notó que los coeficientes son relativamente pequeños. Así, el especialista asoció a esta componente una interpretación de *Infraestructura de Gestión de los RSU, salvo disposición final*. Esta sería, según su apreciación, la componente principal que más impacto tiene en general en el fenómeno de los RSU. De manera similar, en la segunda componente principal contrastó una parte conceptual, de planeación, con respecto a la infraestructura u operación. Así, concluyó que la segunda componente representa la *planeación vs. la operación en la gestión de los RSU*.

Ahora, el especialista generó la *Figura 3*, que corresponde a la puntuación de las dos primeras componentes principales, y en ella él apreció el punto extremo derecho, que corresponde al Distrito Federal, en el cual tuvo el mayor valor. Otros dos valores que destacó son el punto superior, correspondiente a Baja California, con un gran valor en la segunda componente, y Oaxaca con el menor valor en esta última componente principal.

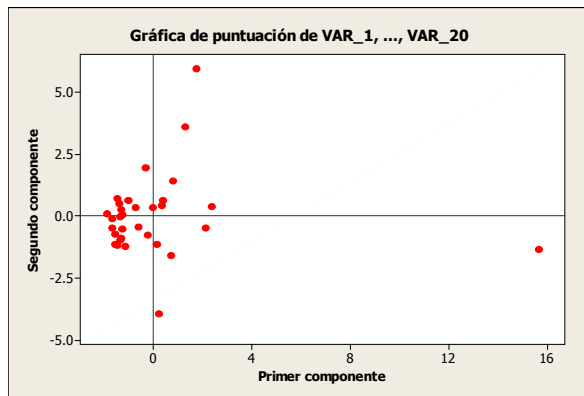


Figura 3. Gráfica de puntuación del ACP sobre datos de RSU.

Fuente: Elaboración del especialista con resultados generados en Minitab®.

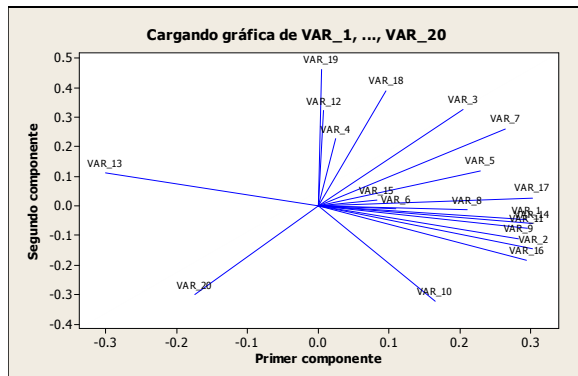


Figura 4. Gráfica de puntuación del ACP sobre datos de RSU.

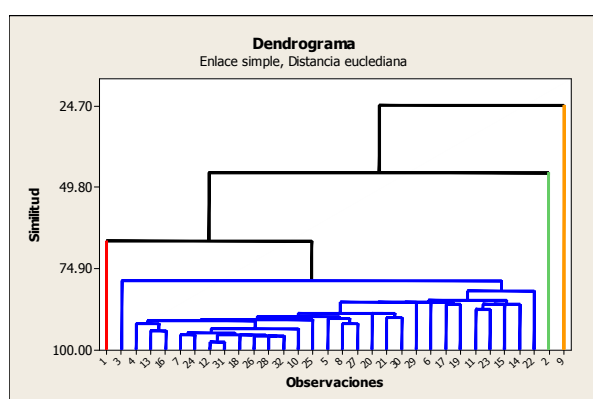
Fuente: Elaboración del especialista con resultados generados en Minitab®.

La figura 4 proporciona información, según el ACP, acerca de la afinidad de las variables, destacando el especialista que “la número 13 parece ser la que marca una *dimensionalidad* distinta con respecto a las restantes, así como la 20 y la 10, aunque éstas en menor medida. Por otro lado, las variables 4, 12 y 19 tienen una gran afinidad en tal sentido”. Así mismo, el especialista muestra el ordenamiento de las Entidades Federativas según la primera componente principal: “Distrito Federal, Baja California, Aguascalientes, Querétaro, Jalisco, Morelos, Estado de México, Nuevo León, Quintana Roo, Guanajuato, Veracruz de Ignacio de la Llave, Oaxaca, Michoacán de Ocampo, Puebla, Tlaxcala, Hidalgo, Colima, Chihuahua, Baja California Sur, Tabasco, Sinaloa, Coahuila de Zaragoza, Chiapas, Tamaulipas, San Luis Potosí, Durango, Sonora, Zacatecas, Campeche, Nayarit, Guerrero y Yucatán. De este ordenamiento observa que las cinco primeras tienen el mayor desarrollo en *infraestructura de gestión de los RSU, salvo disposición final*, mientras que las que tienen mayor área de oportunidad al respecto son las cinco últimas”. Y realizando el ordenamiento con respecto a la segunda componente principal, menciona que los resultados son: “Baja California, Aguascalientes, Distrito Federal, Quintana Roo, Querétaro, Jalisco, Guanajuato, Morelos, estado de México, Colima, Nuevo León, Chihuahua, Coahuila de Zaragoza, Tlaxcala, Baja California Sur, Sinaloa, Tabasco, Campeche, Durango, Hidalgo, Puebla, Michoacán de Ocampo, Nayarit, Sonora, Zacatecas, Veracruz de Ignacio de la Llave, San Luis Potosí, Tamaulipas, Chiapas, Yucatán, Guerrero y Oaxaca”. Observa también que “en la dimensión de la *planeación en*

contraste con la operación en la gestión de los RSU, las cinco entidades federativas que parecen tener mayor desarrollo son las cinco primeras y las menos favorecidas en tal dimensión son las últimas cinco”.

### Empleo del Análisis de Conglomerados

En cuanto al Análisis de Conglomerados, el especialista generó la *Figura 4*, la cual es el *dendrograma* de las entidades federativas bajo consideración, con una distancia euclidiana y con enlace simple (no por grupos), en el cual se distinguieron cuatro conglomerados o colecciones de observaciones. En el primer grupo, de manera aislada, distinguió el Distrito Federal. En el segundo distinguió a Baja California, en el tercero encontró a Aguascalientes y, en el cuarto, todas las demás entidades federativas. Con ello señaló: “respaldo las afirmaciones comentadas en la sección previa del Análisis de Componentes Principales”.



*Figura 4.* Dendrograma de enlace simple y distancia euclidiana de las variables de proporción de RSU.

*Fuente:* Elaboración del especialista con resultados generados en Minitab®.

En la *Figura 5* el especialista mostró cuatro grupos de variables, en las cuales destacó que se dan tres grupos con una sola variable cada uno de ellos: la 20 (*proporción de sitios de disposición final tipo tiradero a cielo abierto reportados como destino de los residuos sólidos urbanos respecto al total en la Entidad Federativa*), la 13 (*proporción de municipios y delegaciones con disponibilidad de servicios de sólo recolección y disposición final de residuos sólidos urbanos respecto al total en la Entidad Federativa*) y la 12 (*proporción del total municipios y delegaciones con disponibilidad de servicios relacionados con los residuos sólidos urbanos respecto al total en la Entidad Federativa*), y el cuarto grupo aglutina a todas las demás. Según sus palabras: “esto es, hasta cierto punto, coincidente con el resultado del ACP, en el cual se contrasta la *infraestructura* (salvo *disposición final de los RSU*), parece marcar una de las diferencias más importantes en este conjunto de datos”.

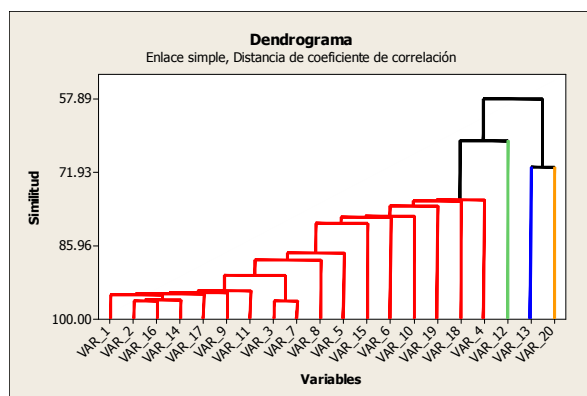


Figura 5. Dendrograma de enlace simple y distancia de coeficiente de correlación de los datos de RSU.

Fuente: Elaboración del especialista con resultados generados en Minitab®

### Conclusiones, reflexiones y comentarios del especialista con respecto a su estudio estadístico

Del resultado de la aplicación del ACP y del AC, el especialista observó que hay dos dimensiones fundamentales en la gestión de la basura: la *Infraestructura de Gestión de los RSU, Salvo Disposición Final*, por un lado, y la *Planeación vs. la Operación en la Gestión de los RSU*, por el otro. El argumento fue, siguiendo textualmente sus palabras: “por supuesto, la designación de estos nombres es finalmente cuestión de decisión personal, y podría ser sujeta a otro tipo de reflexiones por parte de otros investigadores; sin embargo, aquí lo que destaca es que de un fenómeno de 20 dimensiones, dos (que son cruces), resultan ser las que marcan de manera preponderante el fenómeno de los RSU, con lo cual se simplifica conceptualmente su comprensión”.

A su vez, él argumentó que: “el hecho de que las entidades federativas se hubiesen contrastado, bajo aglomeramiento, en básicamente el Distrito Federal, Baja California y Aguascalientes, por un lado, y las demás entidades federativas por el otro nos dice que, a pesar de que estas últimas han dado avances en la gestión de los RSU (básicamente en las tres entidades referidas), existen áreas de oportunidad muy importantes, sobre todo en las entidades del sur del país”.

Finalmente, el especialista aseguró que: “el análisis por supuesto podría no parar aquí y se sería posible aplicar otro tipo de herramientas estadísticas que podrían complementar el mostrado en este documento, aunque creemos que los resultados aquí son representativos y arrojan una gran luz acerca del complejo fenómeno de los RSU”.

### Discusión de resultados

Los siguientes resultados son distinguibles en el estudio llevado a cabo: la selección de un problema para analizar por parte del especialista fue dado por motivos de su adherencia de tal tipo de problemas según su área laboral. Esto es, al estar él inserto a un área de gestión de las estadísticas del medio ambiente valoró como relevante el tomar datos de Residuos Sólidos Urbanos, más que datos de tipo económico, socio-demográfico o de justicia, que son los temas generales que trabaja la institución donde él labora. Esto nos lleva a conjeturar que, en general, los especialistas dentro de sus áreas de trabajo tienden a tomar casos que sean de su interés inmediato y con relevancia a sus labores propias, más que ejemplos en otros contextos. Por

supuesto, esto podría no ser una pauta general y para otros especialistas, que por algún motivo no estuvieran satisfechos en sus actuales áreas, la situación podría ser diferente.

El especialista no optó por la selección de técnicas estadísticas univariadas, como podría ser simplemente la descripción de algunas medidas de tendencia central (media, mediana, moda) o de dispersión (varianza, rango, desviación semi-intercuartílica, etc.), debido a que, según sus palabras, “en fenómenos complejos y de la vida real, estas son técnicas limitadas y más bien deben de tomarse en consideración estrategias multivariadas para el análisis de datos”. Este es un punto extraordinariamente relevante en el cual se distingue un estudiante en algún centro escolar con respecto a los especialistas, los cuales tenderían a tener análisis más elaboradas, y por ser elaborados en sí, sino persiguiendo una descripción más adecuada de la realidad.

Un énfasis muy marcado del especialista fue el perseguir la interpretación de los datos y cifras que se generaban en la aplicación de cada una de las técnicas estadísticas que él utilizó. Esto es, por ejemplo, no se limitó solo a mostrar la varianza explicada de cada componente principal, sino que buscó dar la interpretación contextual de al menos las dos primeras componentes principales, tal y como se mostró en la sección 4.5. Esto, conforme a la experiencia de los autores, marca también de manera muy acentuada una distinción fundamental con respecto a los estudiantes en instituciones de educación superior, los cuales usualmente no les *consterna* en mayor medida la interpretación, sino que se considera que *cumplen* con su tarea estadística al mostrar ciertos conjuntos de cifras. El énfasis en la interpretación muy probablemente se asocia al hecho de que en la vida profesional se busca *explotar* los resultados de manera que dé sentido al proceso de toma de decisiones en contextos y problemas reales del lugar de trabajo.

Con respecto al hecho de seleccionar el Análisis de Componentes Principales y el Análisis de Conglomerados, según la apreciación del especialista, obedece a que son dos de las técnicas de mayor reconocimiento por parte de la comunidad estadística, por un lado, y por el hecho de que sus resultados coadyuvan fuertemente la toma de decisiones y el esclarecimiento conceptual de los problemas. Nuevamente, se observa que la principal inquietud del especialista fue el que los resultados fueran valiosos para el proceso de toma de decisiones y que se pudieran comunicar con relativa facilidad a sus colegas.

Por último, el especialista mostró consuetudinariamente el hábito de poder modificar las *reglas* dadas por los libros para tornarlas más laxas o más exigentes, según fuera el caso, y desde su propia apreciación. Por ejemplo, a pesar de que tomó como mínimo un 70% de la varianza explicada, él se inclinó por considerar que el fenómeno tendría una dimensión de 2, puesto que la diferencia entre el 72.8% y el 61.7% mostrada en la Tabla 1, a pesar de ser reconocida, no le pareció que valiera la pena para volver más complicado el fenómeno.

### Conclusiones

En función a los resultados comentados en la sección previa, son claras las siguientes conclusiones: la forma de aprender y aplicar el aprendizaje de la estadística adquiere una importancia fundamental en el contexto de trabajo y en situaciones reales. En general, esta es una característica muy particular de la Educación Estadística, en la cual el contexto más que ayudar, es indispensable para que adquieran sentido los ejemplos y ejercicios que se desarrollan. En función a lo anterior, parece ser que una buena estrategia que podría emplearse para incentivar el aprendizaje y uso de las técnicas estadísticas es, por un lado, la aplicación de datos reales (y no hacer un sobreuso de los ejemplos académicos que existen en la mayoría de los textos), y por

otro, invitar a especialistas (por ejemplo, de las estadísticas oficiales) a ofrecer sus experiencias estadísticas en grupos académicos.

Finalmente, el contexto en los problemas de Estadística es significativo y es un atributo esencial que no debe ser separado en los ejercicios de Estadística aplicada para que estos en verdad tengan sentido para los estudiantes y cuenten con un segmento interesante y real de interpretación.

### **Limitaciones del estudio**

El presente estudio es un caso de estudio con un egresado de un posgrado especializado en Estadística Oficial. Por esto mismo, entrevemos dos limitantes del estudio: por un lado, sería conveniente llevar a cabo un estudio más amplio con un mayor número de especialistas, egresados del programa de Estadística Oficial, para observar similitudes y contrastar diferencias, y así estar en condiciones de ver si lo observado en el presente estudio es realmente una particularidad, o si es una generalidad; por otro lado, al ser el especialista un egresado de un programa de Estadística, podría existir diferencia contra otros egresados en posgrados distintos, como de Administración o Ingeniería que incluya componentes curriculares estadísticos, por lo que también deben de tomarse con reservar los presentes resultados y no ser generalizables a todo estudiante o egresado de estudios de posgrado en Estadística.

### **Prospectiva**

Los estudios sobre estadística educativa son incipientes en el mundo en general, y particularmente en el caso de México, y más aún aquellos estudios sobre la caracterización de los procesos de enseñanza-aprendizaje en lugares de trabajo y sobre contextos reales y aplicados. Sin embargo, todos ellos son extraordinariamente importantes, tanto en un sentido teórico como práctico, y sin lugar a dudas serán motivo de mayores y más profundos estudios al respecto. En este trabajo se pretendió el brindar un aporte al respecto y esperamos que los resultados sean de interés a los investigadores en el tema y a aquellos que no lo son les parezcan interesantes para poder ser, potencialmente, incorporados en sus propios estudios educativos. Lo que es indiscutible es que las investigaciones seguirán su marcha y dentro de algún tiempo se tengan un cuerpo de conocimiento estructurado y coherente con respecto a este y otros hechos de gran relevancia en la Educación Estadística.

### **Bibliografía y Referencias**

- Bavdaž, M., Ograjenšek, I., & Perviz, L. (2014) Workplace and official statistics: How can higher education contribute to a better relationship? En K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. iase-web.org [© 2014 ISI/IASE]
- Biehler, R. (1994). Probabilistic thinking, statistical reasoning, and the search for causes - Do we need a probabilistic revolution after we have taught data analysis? Versión revisada y extendida de un artículo presentado en el *Fourth International Conference on Teaching Statistics (ICOTS 4), Marrakech, Morocco*, pp. 25-30.
- Brown, J., David, I., Moltchanova, E., Seddon, H., & Harlow, J. (2014) Improving statistical literacy at university. En K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. iase-web.org [© 2014 ISI/IASE]

- Carrasco, J. P. (2014) Education statistics for professional specialization: use of knowledge at the workplace of postgraduate in statistics. En K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. [iase-web.org](http://iase-web.org) [© 2014 ISI/IASE]
- Cobb, G. (2007). Un marco posible para pensar sobre el aprendizaje experimental. *International Statistical Review*, 7(3), 336-347.
- Forbes, S., & Janssen, T. (2014). The use of official statistics in evidence based policy making in new zealand. En K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. [iase-web.org](http://iase-web.org) [© 2014 ISI/IASE]
- Forbes, S., & Keegan, A. (2014) Bringing the workplace into a national certificate in official Statistics. En K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. [iase-web.org](http://iase-web.org) [© 2014 ISI/IASE]
- Instituto Nacional de Estadística y Geografía –INEGI- (2014). Residuos Sólidos Urbanos. En la sección de “Asentamientos y actividades humanas”, sub-sección “Residuos Sólidos Urbanos”. *Módulos ambientales dentro del Censo Nacional de Gobiernos Municipales y Delegacionales 2013*. Recuperado de <http://www3.inegi.org.mx/sistemas/temas/default.aspx?s=est&c=21385>
- Johnson, D. (2000). *Métodos multivariados aplicados al análisis de datos*. México, D.F.: International Thomson Editores.
- Johnson, G., & Shumway, D. (2014). Supporting statistical consultant decision-making within a case-based learning environment- En K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. [iase-web.org](http://iase-web.org) [© 2014 ISI/IASE]
- Johnson, R., & Wichern, D. (2002). *Applied Multivariate Statistical Analysis* (Fifth Edition). New Jersey: United States of America.
- Navarro, J. (2014) Once were warriors: the need of re-education in mathematics and Computing for life "scientisticians". En K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. [iase-web.org](http://iase-web.org) [© 2014 ISI/IASE]
- Moore, D. (1997). New Pedagogy and New Content: The Case of Statistics. *International Statistical Review*, 65(2), pp. 123-165. Recuperado de <http://iase-web.org/documents/intstatreview/97.Moore.pdf>
- Razzouk, N., & Razzouk, J. (2008). Analysis in Teaching with Cases: A revisit to Bloom’s Taxonomy of Learning Objectives. *College Teaching Methods & Styles Journal*, 4(1), 49-56.
- Secretaría de Medio Ambiente y Recursos Naturales –SEMARNAT. (2007). Guía metodológica para el levantamiento y evaluación de información para efectuar el diagnóstico de la gestión integral de los RSU.
- Secretaría de Medio Ambiente y Recursos Naturales –SEMARNAT. (2012). *Informe de la Situación del Medio Ambiente, 2012*. Compendio de Estadísticas Ambientales. Indicadores Clave y de



Desempeño Ambiental. Capítulo 7: Residuos. Recuperado de  
[http://app1.semarnat.gob.mx/dgeia/informe\\_12/pdf/Cap7\\_residuos.pdf](http://app1.semarnat.gob.mx/dgeia/informe_12/pdf/Cap7_residuos.pdf)

- Vaughan, K. (2008). *Workplace learning: A literature review*. Report prepared for Competenz.
- Wild, C., & Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review*, 67(3), 223-265.
- Westbrooke, I., & Ellis, P. (2014). Training to develop modern statistics in the workplace using R and R Commander - Experiences from the New Zealand government sector. En K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. [iase-web.org](http://iase-web.org) [© 2014 ISI/IASE]