

Razonamiento Estadístico de Estudiantes Universitarios sobre el Análisis de Datos en un Ambiente Computacional

Statistical Reasoning of University Students on Data Analysis in Computational Environment

Santiago Inzunza Cazares*

Resumen

En el presente artículo se presentan resultados de una investigación con una muestra de 34 estudiantes universitarios (18-19 años) acerca del razonamiento estadístico que desarrollaron al analizar datos mediante un enfoque exploratorio en un ambiente computacional caracterizado por la disponibilidad de diversas representaciones dinámicas e interactivas (gráficas, numéricas y simbólicas). Los resultados muestran que en las primeras actividades los estudiantes se plantearon preguntas y analizaron los datos involucrando, principalmente, un valor o un caso individual dentro de un grupo de datos; es decir, mostraron una comprensión local de los datos. Sin embargo, al final del estudio lograron enfocarse - aunque de forma insuficiente - en el reconocimiento de patrones y tendencias en los datos, desarrollando con ello una comprensión más global de los datos. Se observó una tendencia a depender más de tablas que de gráficas, haciendo mayor énfasis en promedios (media o mediana), dada la familiaridad que los estudiantes tenían con estos conceptos de sus estudios previos.

Palabras Clave: Análisis de datos. *Software* Estadístico. Razonamiento Estadístico.

Abstract

In this article, research results are presented for a sample of 34 university students (18-19 years) about statistical reasoning developed to analyze data through an exploratory approach in a computational environment characterized by the availability of various dynamic and interactive representations (graphical, numerical, and symbolic). The results show that, in the first activities, the students proposed questions and analyzed the data, primarily engaging a value or an individual within a group of data; that is, they exhibited a local data understanding. However, at the end of the study they were able to focus – insufficiently -- in recognizing patterns and trends in data, developing a more comprehensive data understanding. We observed a tendency to depend more on tables than on graphics, putting more emphasis on averages (mean or median), given the familiarity that students have with these concepts in their previous studies.

Keywords: Data analysis. Statistical Software. Statistical Reasoning.

* Doctor en Ciencias en la especialidad de Matemática Educativa por el Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN). Profesor e Investigador en la Facultad de Informática (Campus Culiacán) de la Universidad Autónoma de Sinaloa (UAS), Culiacán, Sinaloa, México. Dirección Postal: Circuito de las Amapas, 785-51, Residencial Amapas, CP 80060, Culiacán, Sinaloa, México. E-mail: sinzunza@uas.edu.mx.

1 Introducción

Recientes reformas curriculares para la enseñanza de la estadística, en muchos países, (por ejemplo: NCTM, 2000; MEC, 2006; ME, 2007) otorgan especial importancia al desarrollo de competencias tales como el razonamiento estadístico, el pensamiento estadístico y la resolución de problemas. Algunos aspectos de enseñanza que se destacan en dichas reformas son la promoción de un aprendizaje activo, uso de más datos (reales y en contexto) y conceptos, menos fórmulas y procedimientos, y automatización de cálculos y gráficas hasta donde sea posible mediante el uso de tecnología. Estos cambios en contenido y pedagogía han dado lugar por un lado, al desarrollo de nuevas herramientas de *software* que hacen posible ambientes de aprendizaje más exploratorios, como es el caso de Fathom (FINZER, 2002) y TinkerPlots (KONOLD; MILLER, 2005) por mencionar algunas. Por otra parte, en el aspecto pedagógico, se han propuesto modelos de enseñanza que promueven el razonamiento y el pensamiento estadístico a través de ambientes constructivistas en los cuales se contempla el uso de herramientas tecnológicas (por ejemplo, COBB; MCCLAIN, 2004; GARFIELD; BEN-ZVI, 2008).

En este sentido Lajoie y Romberg (1998) proponen una agenda para la enseñanza y evaluación de la estadística e identifican como una tarea importante para el desarrollo del razonamiento y pensamiento estadístico de los estudiantes de bachillerato, formular preguntas de investigación y responderlas mediante el análisis de los datos; señalan, a su vez, que se debe dar importancia al planteamiento del problema y al proceso de planeación y recolección de datos. Wild y Pfannkuch (1999) definen el ciclo investigativo de un problema (Problema, Plan, Datos, Análisis y Conclusiones) como parte de un modelo para el desarrollo de pensamiento estadístico de los estudiantes. Burrill et al. (2003), por su parte, distinguen las siguientes competencias como fundamentales para desarrollar el razonamiento estadístico:

1. Formular preguntas que puedan ser abordadas con datos, recolectarlos, organizarlos y desplegar aspectos relevantes para responderlas.
2. Seleccionar y utilizar métodos estadísticos apropiados para analizar los datos.

La Asociación Americana de Estadística (ASA por sus siglas en inglés) a través del proyecto GAISE (FRANKLIN et al., 2005) establece que la enseñanza de la estadística se debe enfocar en resolución de problemas estadísticos, usando datos reales, aprendizaje activo y uso apropiado de tecnología. En el proyecto se define un proceso de cuatro pasos para resolver un problema estadístico: formular preguntas que pueden ser abordadas a través de la recolección y análisis de los datos, recolectar y organizar los datos en forma apropiada,

analizar los datos e interpretar resultados. De igual forma, el currículo de estadística del sistema escolar de Nueva Zelanda (ME, 2007, p.26) para los niveles preuniversitarios establece:

[...] la estadística involucra identificar problemas que pueden ser explorados por el uso de datos apropiados, diseño de investigaciones, recolección de datos, explorar y utilizar patrones y relaciones en los datos, resolver problemas y comunicar resultados. La estadística involucra además, interpretar información estadística, evaluar argumentos basados en datos y tratar con la incertidumbre y la variación.

De esta manera, en muchos currículos se propone que el análisis haga más énfasis en la exploración que en la descripción de los datos; en dicha exploración, las representaciones gráficas ocupan un lugar fundamental. Este cambio de enfoque ha sido derivado del trabajo de John Tukey (1977) y se conoce como *análisis exploratorio de datos*. Una idea fundamental en la que se apoya este enfoque es que al utilizar diferentes representaciones de un conjunto de datos se puede facilitar la comprensión. En un ambiente computacional dinámico e interactivo es posible cambiar de una representación a otra (por ejemplo, de una tabla a una gráfica o de una gráfica a otra gráfica) de una forma sencilla, con lo que se facilita la visualización de estructuras en los datos; también es posible identificar patrones en los datos mediante el cambio de un dato, un parámetro o la escala de una gráfica.

De acuerdo con lo anterior, el análisis exploratorio de datos puede ser una herramienta de utilidad en la generación de hipótesis, conjeturas y preguntas de investigación acerca de los fenómenos de donde los datos fueron obtenidos, y por ello se propone como medio para desarrollar una comprensión global de los datos y promover el desarrollo del razonamiento estadístico en los estudiantes.

2 Antecedentes

En los últimos 15 años han empezado a emerger, cada vez más, propuestas y estudios de investigación en los cuales se analiza la influencia de la tecnología computacional sobre diversos aspectos relacionados con el desarrollo del razonamiento y pensamiento estadístico que se involucra en el análisis exploratorio de datos. Por ejemplo, Biehler (1997) reporta resultados de dos estudios realizados para identificar áreas problemáticas y dificultades que tienen estudiantes de bachillerato en un curso elemental de análisis de datos basado en proyectos con el uso del *software* DataScope. En dicho estudio, señala que los estudiantes superaron el obstáculo consistente en que un problema de análisis de datos tiene un resultado único, pero tuvieron dificultades para interpretar y describir verbalmente algunas gráficas y

tablas. Mostraron, además, un pensamiento centrado en promedios al comparar distribuciones, dejando de lado otras propiedades importantes de las distribuciones. Al interpretar los datos muchos estudiantes tuvieron dificultades para distinguir entre el conjunto y el nivel individual de los datos.

Ben-Zvi (2002) realizó un estudio con estudiantes israelíes de secundaria (13 años) para evaluar su habilidad de dar sentido a los datos y a sus representaciones en un ambiente exploratorio con una hoja de cálculo mediante el enfoque de plantear preguntas y hacer hipótesis sobre un conjunto de datos reales proporcionado previamente. Los resultados muestran que muchos estudiantes se enfocaron en propiedades de casos individuales (por ejemplo, un punto de la gráfica o un caso de una tabla) en lugar de enfocarse en propiedades globales de los datos como conjunto (tendencia, centralidad, variabilidad). Esto sucedió particularmente en los casos que requerían preguntas abiertas, sin embargo, cuando las preguntas fueron más cerradas se incrementó la visión global de las propiedades de los datos. Muchos estudiantes también planteaban preguntas e interpretaban resultados, tomando en cuenta el contexto de los datos en lugar de fijarse en sus propiedades estadísticas. Sin embargo, como resultado final se observó que los estudiantes dieron sentido correcto a diversos aspectos de los datos desde el punto de vista de un experto.

Por su parte, Bakker, Derry y Konold (2006) diseñaron un experimento usando el *software* TinkerPlots con el propósito de mejorar las habilidades de exploración de datos de estudiantes de 11 años. Los autores concluyen que el trabajo con proyectos estadísticos y *software* puede ayudar a los estudiantes a desarrollar lenguaje estadístico, a familiarizarse con las gráficas para extraer información de ellas y empezar a pensar intuitivamente con conceptos estadísticos tales como las medidas de tendencia central y la dispersión.

Monteiro et. al. (2010), utilizando el mismo *software* (TinkerPlots) realizaron un trabajo con estudiantes brasileños próximos a graduarse de profesores de primaria, con el propósito de explorar los límites y posibilidades que tiene la herramienta para comprender las relaciones entre datos. Los resultados proporcionan argumentos para apoyar la introducción de situaciones basadas en computadora para enseñar nociones estadísticas en el currículum de la formación de profesores de primaria.

En este mismo sentido, Vizcarra e Inzunza (2011) estudiaron el efecto que tienen el uso de actividades didácticas y el *software* TinkerPlots en el desarrollo del razonamiento estadístico de estudiantes mexicanos de tercer grado de bachillerato (17-18 años). Los resultados señalan que las actividades con el apoyo del *software* tuvieron un efecto positivo en el desarrollo del razonamiento estadístico de los estudiantes. Las respuestas

proporcionadas a los diferentes ítems se ubican en un nivel cognitivo alto (nivel relacional) de la taxonomía SOLO (BIGGS; COLLIS, 1982). Como consecuencia los resultados del estudio, sugieren enfocar más la enseñanza de la estadística hacia el desarrollo de procesos de mayor nivel cognitivo que el simple manejo de fórmulas y procedimientos.

Utilizando otra herramienta computacional (Fathom), Inzunza y Juárez (2010), reportan un estudio con profesores de bachillerato mexicanos, acerca del razonamiento estadístico que desarrollaron al analizar datos en un ambiente computacional. Sus resultados señalan que, a pesar de la abundancia de representaciones proporcionadas por el *software*, los profesores se enfocaban en el cálculo de promedios principalmente para describir y comparar distribuciones de datos, y, en menor medida, utilizaban otras propiedades de los datos tales como la forma y la variabilidad. Muchos profesores construyeron interesantes gráficas que reflejaban importantes propiedades de los datos, pero no se apoyaron en ellas para el análisis, dado su poco conocimiento estadístico.

Dado que estudios previos se han enfocado principalmente en estudiantes y profesores de niveles preuniversitarios, en el presente trabajo nos hemos propuesto investigar en torno al potencial de un ambiente exploratorio de datos con un *software* de estadística dinámica diseñado con propósito de enseñanza, en el desarrollo de una comprensión global de los datos y razonamiento estadístico de estudiantes universitarios. En particular, nos hemos planteado las siguientes preguntas: ¿qué tipo de representaciones y recursos del *software* utilizan los estudiantes para responder preguntas acerca de un conjunto de datos?, ¿cómo se caracteriza el razonamiento estadístico desarrollado por los estudiantes en el enfoque exploratorio de los datos mediante el uso de tecnología?, ¿utilizan todo el potencial de las representaciones que proporciona el *software* para el análisis de los datos para el desarrollo de una comprensión más global de los datos?.

3 Marco teórico

La perspectiva teórica que adoptamos en el presente trabajo tiene que ver, en primer lugar, con la visión de un experto estadístico que requiere plantearse preguntas acerca de un conjunto de datos, poner en juego conocimientos y habilidades para dar sentido a las complejas relaciones e información que expresan algunas gráficas y tablas descriptivas con resúmenes estadísticos para interpretar los resultados y emitir conclusiones. Esta perspectiva ubica al razonamiento y el pensamiento estadístico como competencias centrales en el proceso de aprendizaje del análisis de datos. En particular, Garfield (2002) señala que el

razonamiento estadístico puede ser definido como la manera en la cual las personas razonan con ideas estadísticas y el sentido que le dan a la información estadística, lo cual implica hacer interpretaciones basadas en conjuntos de datos y sus representaciones.

El tipo de representaciones utilizadas en el análisis de datos (tablas, gráficas, fórmulas) incorporan elementos que resultan complejos de interpretar por muchos estudiantes. De acuerdo con Ben-Zvi y Arcavi (2001), el análisis requiere poner en juego visiones locales y globales de los datos. Una comprensión local de los datos significa enfocarse en algún valor individual (o en algunos) dentro de un conjunto de datos (por ejemplo, la entrada de una tabla o un punto de una gráfica). La comprensión global requiere de habilidades para buscar, reconocer, describir y explicar patrones generales en un conjunto de datos (por ejemplo, tendencias, variabilidad, forma) ya sea a través de la visualización o por medio del cálculo de medidas estadísticas.

En segundo lugar, el empleo de tecnología computacional en el análisis de datos no solo representa un recurso *amplificador* de la capacidad de cálculo y construcción de representaciones gráficas - aspectos de suma importancia en el trabajo estadístico -, sino que constituye, a su vez, un recurso con potencial *reorganizador* de la mente de los usuarios, con capacidad para provocar cambios estructurales en el sistema cognitivo a través de la organización y transformación de las diversas representaciones de los datos y la visualización de patrones que emergen al realizar cambios en algunas de las componentes de dichas representaciones o en los datos mismos. Lo anterior le otorga a la tecnología computacional un *status* de herramienta cognitiva en el sentido establecido por Pea (1987, p.91), quien la define como “cualquier medio que ayuda a trascender las limitaciones de la mente, en el pensamiento, el aprendizaje y las actividades de resolución de problemas”.

Pea (1987) señala que una característica común de las herramientas cognitivas es que vuelven externos los productos intermedios del pensamiento, los cuales pueden ser analizados y discutidos. Particularmente, en el caso de las computadoras, constituyen una extraordinaria y potente herramienta cognitiva para aprender a pensar matemáticamente; con ellas se pueden operar no sólo números, sino también símbolos, y permiten almacenarlos y manipularlos dinámicamente, y permiten interacciones con los usuarios en tiempo real.

4 Metodología

En la investigación participaron 34 alumnos (18-19 años) de primer grado de la carrera de Informática mientras tomaban el curso de Probabilidad y Estadística. Los conocimientos

estadísticos de los sujetos no fueron evaluados previamente; sin embargo, otros estudios recientes que se han realizado con estudiantes similares que cursan el último grado de bachillerato (INZUNSA; VIZCARRA, 2009) han mostrado un conocimiento superficial de los conceptos estadísticos que forman parte de un curso básico de análisis de datos como el que tomaron los estudiantes participantes en el estudio.

La investigación se llevó a cabo mientras se abordaban los temas de análisis de datos que se contemplan en el programa de estudios (recolección de datos, representaciones de los datos, tendencia central, variabilidad y correlación), y se enfocó hacia el uso de una perspectiva exploratoria con el *software* dinámico Fathom. En cada una de las tres actividades se proporcionó un archivo de datos con diversas variables. En la primera actividad se solicitaba a los estudiantes que plantearan preguntas y, posteriormente, realizaran el análisis correspondiente para responderlas; sin embargo, en las siguientes dos actividades, se proporcionaron las preguntas por el profesor-investigador, dada la dificultad que tuvieron para su planteamiento en la primera actividad. Los estudiantes reportaron los análisis estadísticos (tablas, gráficas y medidas descriptivas) en archivos de computadora y utilizaron la opción de texto que el mismo *software* Fathom proporciona para describir sus interpretaciones y las respuestas a las preguntas planteadas.

Al momento de las actividades con el *software*, los estudiantes ya habían visto los temas del apartado de análisis de datos descritos anteriormente y habían aprendido el uso básico del *software*, por lo que asumimos que tenían los conocimientos estadísticos y del uso de la herramienta para realizar el análisis de los datos que les fueron presentados. Los archivos de datos contemplaban tanto variables cualitativas como variables cuantitativas, por lo que las preguntas involucraban variables cualitativas y cuantitativas en forma separada, análisis de una variable cualitativa contra una cuantitativa y dos variables cuantitativas de manera conjunta.

Para el análisis de los resultados de cada actividad se hizo un análisis *a priori* sobre las representaciones y el tipo de análisis que desde el punto de vista de un experto estadístico se podría realizar para responder las preguntas planteadas. Se consideraron elementos de naturaleza cuantitativa, que permitieron elaborar tablas con la frecuencia que los estudiantes utilizaron cada representación. A su vez, utilizamos elementos de naturaleza cualitativa (análisis de respuestas escritas a cada actividad) para proporcionar descripciones de las interpretaciones más representativas que los estudiantes hicieron de los datos (señaladas en letras *itálicas*), resaltando aspectos correctos del análisis y señalando dificultades y errores en el manejo de los conceptos involucrados. En este sentido, las interpretaciones seleccionadas

en la discusión de los resultados corresponden a aspectos del análisis – correctos e incorrectos – que consideramos importantes de resaltar.

5 Resultados y discusión

El reporte de resultados contempla dos de las tres actividades desarrolladas durante el curso. La primera actividad tuvo un carácter introductorio y, en ella, se planteó que los estudiantes formularan preguntas estadísticas para responderlas con los datos proporcionados, en específico se les solicitó el planteamiento de tres preguntas que contemplaran cualquiera de las 11 variables (5 cualitativas y 6 cuantitativas) de un conjunto de 63 computadoras que se venden en el mercado mexicano, cuya información fue tomada de la Revista del Consumidor (No. 402, 2010). Una parte del archivo de datos se muestra en el Cuadro 1.

Computadoras					
	Tipo	Marca	Modelo	Sistema_Operativo	Microprocesador
1	Computadora Portátil Laptop	Toshiba	Satellite A505-SP6023	Windows 7 Home ...	Intel Core i3 M330
2	Computadora de escritorio...	Gateway	SX2840-10m	Windows 7 Home ...	Intel Core i5 650
3	Computadora de escritorio...	Lanix	Magnum GX4070	Windows 7 Ultimat...	Intel Core i7 920
4	Computadora de escritorio...	HP	Pavillion PCP6380LA	Windows 7 Home ...	Intel Core i5 650
5	Computadora Portátil Laptop	Toshiba	Qosmio X505-SP8017M	Windows 7 Home ...	Intel Core i7 Q720
6	Computadora Portátil Laptop	Sony	Vaio VPCF110FL/B	Windows 7 Home ...	Intel Core i7 Q720
7	Computadora Portátil Laptop	HP	Pavillion DV6-2187LA	Windows 7 Home ...	Intel Core i7 Q720
8	Computadora Portátil Laptop	HP	Pavillion DV6-2177LA	Windows 7 Home ...	AMD Turion II Ultra...
9	Computadora Portátil Laptop	Toshiba	Satellite L505D-SP601	Windows 7 Home ...	AMD Turion II X2 M...
10	Computadora Portátil Laptop	Gateway	NV5901E	Windows 7 Home ...	Intel Core i5 M430

Velocidad...	Capacidad...	Memoria_RAM_GB	Tamaño...	Peso...	Precio_...
2.13	500	6	16	2.9	15999
3.2	1000	6	20	0	15999
2.67	1000	6	23.5	0	31929
3.2	1000	6	25	0	22989
1.6	500	4	18.4	4.7	32499
1.6	500	4	16.4	3.1	29499
1.6	500	4	15.6	2.8	18999
2.5	500	4	15.6	2.8	16499
2.3	500	4	15.6	2.8	14375
2.27	320	4	15.6	2.7	15999

Cuadro 1 - Datos sobre diversas variables de un conjunto de computadoras que se venden en México

El análisis revela que la mayoría de los estudiantes tuvo dificultades para plantear preguntas que permitieran conocer el comportamiento global del conjunto de datos (por ejemplo, tendencia central, variabilidad, forma de la distribución). En cambio se plantearon preguntas que involucraban un solo dato (máximo o mínimo, generalmente), una marca o una determinada computadora; solo 2 de los 34 estudiantes relacionaran más de una variable. Las preguntas sobre variables numéricas fueron más frecuentes. Ejemplo de las preguntas planteadas se muestran a continuación:

- ¿Qué marca de computadoras tiene mayor memoria RAM?
- ¿Qué tipo de computadoras es la de menor peso?
- ¿Cuántas computadoras tiene sistema operativo Windows?
- ¿Cuántas computadoras tiene procesador Intel I7?
- ¿Cuáles son las computadoras más económicas?
- ¿Cuál es la computadora con el procesador más veloz?

Dado que las preguntas acerca del comportamiento de un conjunto de datos constituyen un indicador del nivel de razonamiento acerca de los datos y la comprensión de los elementos que los describen, podemos señalar que los estudiantes mostraron una comprensión muy local al enfocarse en datos individuales o entradas aisladas de las tablas de datos, en lugar de enfocarse en propiedades globales de los datos vistos como un conjunto. Estudios realizados por Biehler (1997) y Ben-Zvi (2002) reportan resultados similares.

La tercera actividad (final) trataba de un conjunto de datos sobre automóviles que se venden en México y fueron recopilados del periódico Reforma (REFORMA, 2010). Las preguntas fueron planteadas por el investigador y tenían el propósito de evaluar diferentes propiedades de los datos mediante el uso de representaciones gráficas y numéricas. La actividad se describe a continuación:

Con los datos del archivo automóviles, responder lo que se indica, construyendo las gráficas y realizando los cálculos que consideres necesarios para ello. Realiza el análisis de la forma más completa posible.

1. *Analiza la variable marca de automóviles e interpreta los resultados.*
2. *Analiza la variable precio e interpreta los resultados.*
3. *Analiza la variable Precio contra Tipo de Transmisión e interpreta los resultados.*
4. *Analiza la variable precio contra potencia (HP) e interpreta los resultados.*
5. *¿Cuáles son las tres marcas más económicas (en orden ascendente). Explica porqué.*

Automóviles							
	Marca	Tipo	Precio	Motor	HP	Tipo_de_Transmisión	<n
62	Nissan	Pathfinder	355	V6	240	Automatico	
63	Nissan	Murano	359	V6	235	Automatico	
64	Honda	Odyssey	365	V6	240	Automatico	
65	Chevrolet	Sonora	380	V8	295	Automatico	
66	Ford	Expedition	391	V8	260	Automatico	
67	Nissan	Maxima	396	V6	265	Automatico	
68	Honda	Pilot	399	V6	240	Automatico	
69	Chevrolet	Suburban	417	V8	295	Automatico	
70	Ford	Thunder...	499	V8	280	Automatico	

Cuadro 2 - Datos y precios (miles de pesos) de vehículos que se venden en México

Un análisis *a priori* de las preguntas nos permite anticipar el tipo de análisis y las representaciones permitidas por el *software*, que desde la visión de un experto podrían utilizar los estudiantes para responderlas (ver Cuadro 3).

Representación posible de utilizar	Preguntas				
	1	2	3	4	5
Diagrama de barras	X				
Diagrama de cajas		X	X		X
Diagrama de puntos		X	X		X
Diagrama de barras acumuladas	X				
Histograma		X	X		
Tabla resumen con medidas descriptivas	X	X	X	X	X
Diagrama de dispersión				X	
Alguna gráfica y tabla con medias descriptivas	X	X	X	X	

Cuadro 3 - Representaciones posibles de utilizar para responder las preguntas sobre los datos de los automóviles

Análisis de la pregunta 1

La pregunta involucra una sola variable cualitativa y corresponde al nivel más sencillo de análisis de datos. Los estudiantes utilizaron las representaciones previstas para responder la pregunta, y en algunos casos, incluso, utilizaron más de una representación (ver Cuadro 4). Un ejemplo del uso de dos representaciones (gráfica y tabla) en forma simultánea se muestra en la figura 1.

Representaciones	Cantidad
Diagrama de barras	32
Tabla de frecuencias	16
Diagrama de barras apiladas	1
Diagrama de barras y tabla de frecuencias	15

Cuadro 4 - Representaciones utilizadas para el análisis de la variable *marca de automóvil*

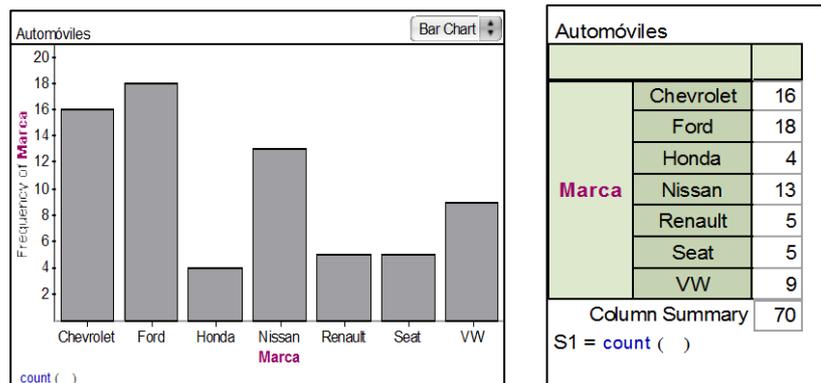


Figura 1 - Diagrama de barras y tabla de frecuencias construidos por Bianca para el análisis de la variable marca de automóvil

No obstante que el diagrama de barras acumuladas (diagramas de listón) permite presentar la información en forma porcentual, solo un estudiante lo utilizó (ver figura 2), la mayoría optaron por el diagrama de barras que era una representación más conocida por ellos.

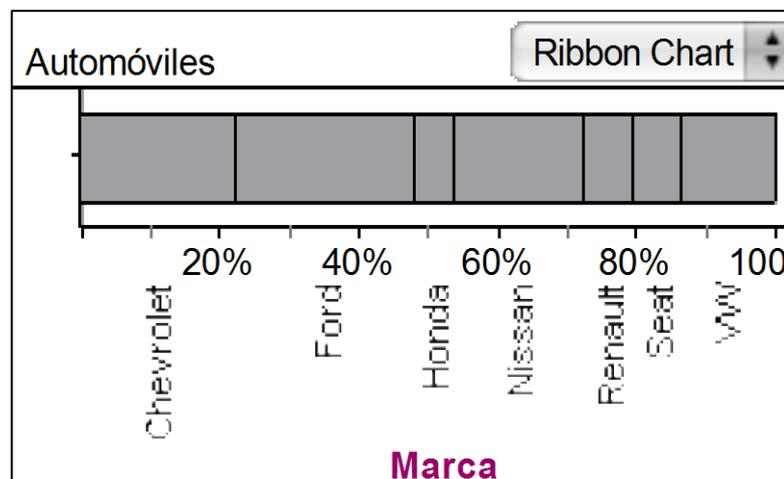


Figura 2 - Diagrama de barras apiladas construido por Geovanny

El análisis de las interpretaciones que hicieron los estudiantes acerca de los datos muestra que no siempre obtuvieron toda la información posible que las representaciones les proporcionaban. En la mayoría de los casos se concretaron a describir la categoría más alta o las más pequeña, o bien, las marcas de automóviles que ocupan las primeras o últimas posiciones en ventas. Solo dos estudiantes hicieron comparaciones entre categorías, situación

que estaba prevista en el análisis a priori pero que casi no fue utilizada. Un ejemplo de ello se muestra a continuación:

Ford es la marca dominante con 18 automóviles en la tabla, superando así con el cuádruple a la menor que es Honda con 4 automóviles.

Análisis de la pregunta 2:

Del Cuadro 5 se observa que 24 estudiantes utilizaron más de una representación (gráficas y tablas) para responder la pregunta planteada. Entre quienes construyeron tablas hubo variación en cuanto a las medidas descriptivas que utilizaron: 15 solo calcularon la media y/o mediana, 7 calcularon las cinco medidas resumen (mediana, mínimo, máximo, primer cuartil y tercer cuartil) y 2 utilizaron la opción de medidas básicas que ofrece el software (media, desviación estándar, conteo, error estándar y datos faltantes).

Representaciones utilizadas	Cantidad
Diagrama de puntos	11
Histograma	9
Diagrama de caja	3
Tabla resumen con medidas descriptivas	25
Gráficas y tabla con medidas descriptivas	24
Gráficas incorrectas	11

Cuadro 5 - Representaciones utilizadas para el análisis de la variable *Precio del Automóvil*

Sin embargo, el análisis de las interpretaciones de los datos revela que las representaciones construidas fueron utilizadas en forma superficial para extraer información acerca de los datos; además, en muchos casos las interpretaciones se limitaron a mencionar valores numéricos obtenidos por el *software*. Por ejemplo, Bianca construye una diagrama de puntos y coloca marcas que identifican la media y la mediana, además construye una tabla con la media, mediana y valor mínimo. Se apoya en ambas representaciones para realizar el análisis de los datos, pero no extrae toda la información que le proporciona el diagrama de puntos (por ejemplo, variabilidad, forma, valores extremos).

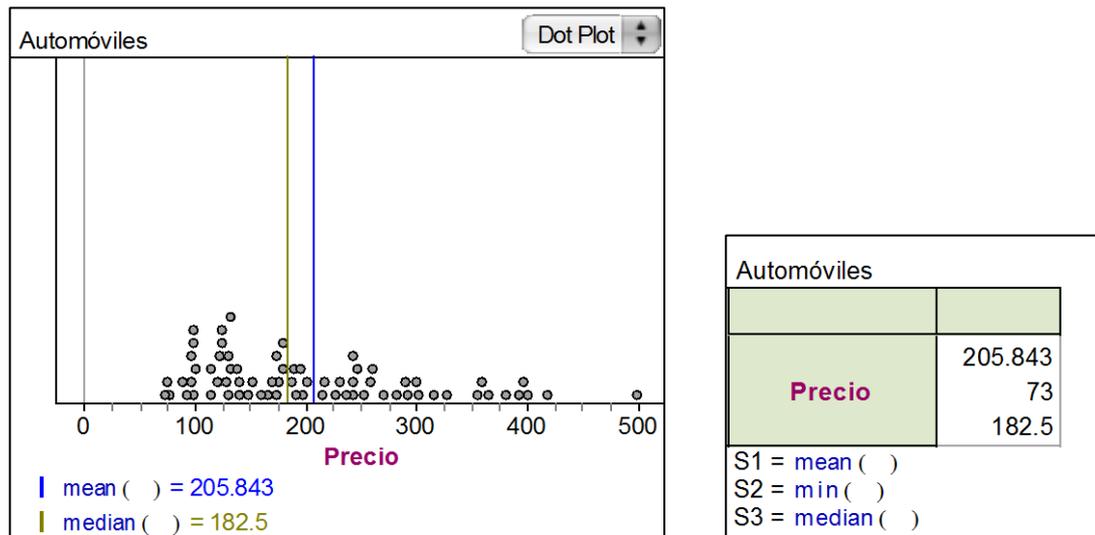


Figura 3 - Representaciones de datos construidas por Bianca

La media aritmética de los precios es de 205.8 miles de pesos, mientras que la mediana equivale a 182.5 miles de pesos. El precio más bajo que aparece es de 73 mil pesos, mientras que el más alto es de 499 mil pesos. En la gráfica también se observa que aproximadamente el 60% de los precios se encuentran por debajo de la media de 205.8

Daniel utiliza un diagrama de puntos, un diagrama de caja y una tabla con las 5 medidas resumen de un diagrama de caja.

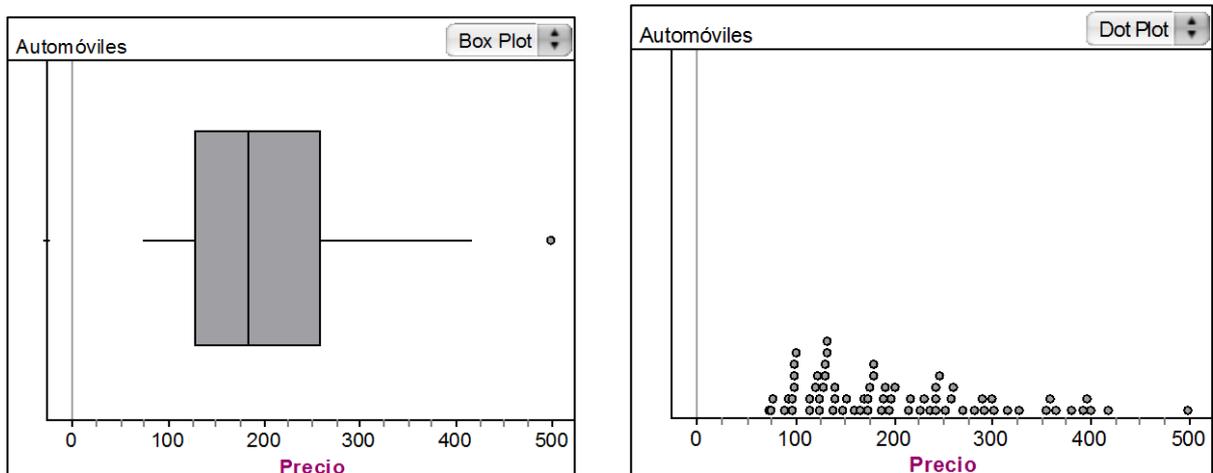


Figura 4 - Representaciones de datos construidas por Daniel

Cerca del 75% de los automóviles se encuentran por debajo de los 260 mil pesos. El automóvil con el mayor precio cuesta 499 mil, el automóvil con el menor precio es de 73 mil. La media de los precios es de \$205.8 mil y la mediana es de \$182.5 mil.

Se observa que Daniel utiliza correctamente el tercer cuartil del diagrama de caja para realizar su interpretación, utiliza la tabla para mencionar el valor máximo, el valor mínimo y los dos promedios; sin embargo, el diagrama de puntos que construyó no lo utiliza en la descripción. Geovanny, por su parte, construye un diagrama de caja y una tabla con las cinco medidas resumen que proporciona el mismo diagrama.

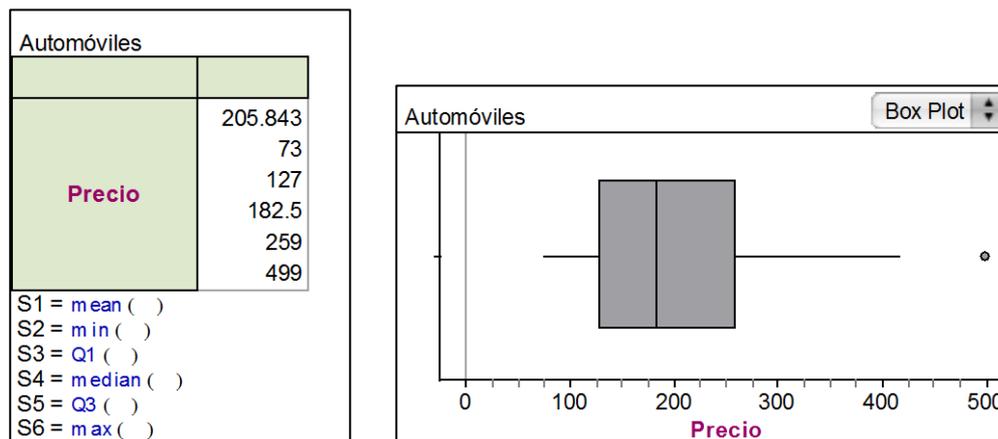


Figura 5 - Representaciones de datos construidas por Geovanny

El precio más alto es de 499 mil pesos y el precio más bajo es de 73 mil pesos. La media de los precios es de 205.8 mil pesos. El primer cuartil nos representa precios menores a 127 mil pesos. El segundo cuartil abarca de 127 mil pesos a 182.5 mil pesos. El tercer cuartil va desde 182.5 hasta 259 mil pesos y el último cuartil va desde 259 hasta 417 mil pesos. La mediana es de 182.5

Se observa que Geovanny se limita a describir los resultados que le proporciona el *software*, pero no recurre a una interpretación que muestre una comprensión del comportamiento de los datos a través de dichas medidas. En este mismo sentido vemos la interpretación de María, que utiliza las mismas representaciones con algunos elementos adicionales:

Se puede ver fácilmente que más del 50% del total de los automóviles tienen un precio que está entre los 90,000 y los 250,000 pesos. La media aritmética del precio de los automóviles es de 205.8. El precio más bajo es de 73,000 pesos y el más alto es de 499,000 pesos (ambos son de la marca Ford)

Análisis de la Pregunta 3

Esta pregunta involucra una variable cualitativa y una cuantitativa, por lo que el análisis implica realizar comparaciones entre las dos distribuciones de los precios que se generan al considerar el tipo de transmisión (manual y automática). Los estudiantes utilizaron las representaciones que se describen en el siguiente cuadro:

Representaciones utilizadas	Cantidad
Diagrama de puntos doble	15
Diagrama de caja doble	7
Histograma doble	4
Tabla resumen con medidas descriptivas y alguna gráfica	22
Tabla resumen con medidas descriptivas solamente	1
Gráficas incorrectas	6

Cuadro 6 - Representaciones utilizadas para el análisis de las variables *precio* y *tipo de transmisión*

De nuevo la mayoría de los estudiantes recurre a más de una representación para analizar los datos. Veamos, primeramente, ejemplos de quienes construyeron solo una gráfica.

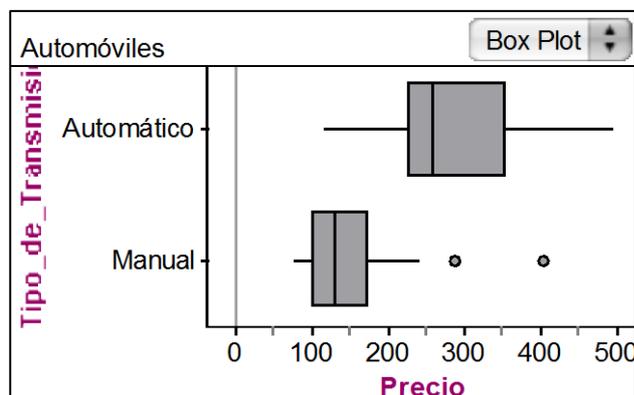


Figura 6 - Gráfica de caja construida por Manuel

Se puede ver que el tipo de transmisión manual es de menor precio que la automática, ya que la mediana está en un precio más bajo que la de transmisión automática. El de menor precio de transmisión manual es de 75 mil mientras que el de automático es de 113 mil.

Manuel interpreta bien el diagrama de cajas y utiliza la tabla original con los datos para identificar el valor mínimo y máximo, que pudo haber calculado mediante una tabla de medidas descriptivas. Otro estudiante (René) utilizó también solo el diagrama de caja y abunda más en su descripción:

Se puede observar que los autos automáticos tienen mayor variabilidad en los precios. La media en los autos de transmisión automática se localiza entre los 250 y 260, en cambio la media de los autos de transmisión manual se localiza entre los 100, 150, lo cual puede resumir que los autos de transmisión automática basándonos en los datos de media son casi o el doble más caros que los de transmisión manual.

René ha obtenido mayor información de la gráfica al visualizar variabilidad, centro y realiza una comparación correcta al establecer que el precio de los automáticos es del doble de los manuales. Comete el error de confundir la mediana con la media. Atanael, por su parte, corresponde al grupo de los alumnos que, además del diagrama de cajas, elaboró una tabla con medidas descriptivas.

Automóviles		Precio
Tipo_de_Transmisión	Automático	33 280.606
	Manual	37 139.162
Column Summary		70 205.843

S1 = count ()
 S2 = mean ()

Figura 7 - Tabla con medidas descriptivas construida por Atanael

La gráfica nos muestra que la transmisión automática es más cara que la manual en la mayoría de los casos, teniendo la automática una media de 280.6 y la manual una media de 139.1

Puede verse que el alumno complementa las dos representaciones, el diagrama de caja le permite realizar la aseveración: *la transmisión automática es más cara que la manual en la mayoría de los casos*, mientras que la tabla le permite precisar los valores de la media. Comete el error de la media.

Alejandro construye un diagrama de caja y obtiene alguna información correcta de la gráfica cuando señala el porcentaje de automóviles automáticos que son más caros que los manuales. No utiliza la tabla para el análisis, aún cuando la ha construido. Comete error al referirse a la media en lugar de la mediana.

Según la gráfica de caja, por encima del 50% de los automóviles automáticos son más caros que casi todos los manuales y que mientras que la media de los automóviles manuales se encuentra por debajo de los 150 de precio. La media de los automáticos está cerca de los 300.

Por su parte Bianca complementa el diagrama con una tabla:

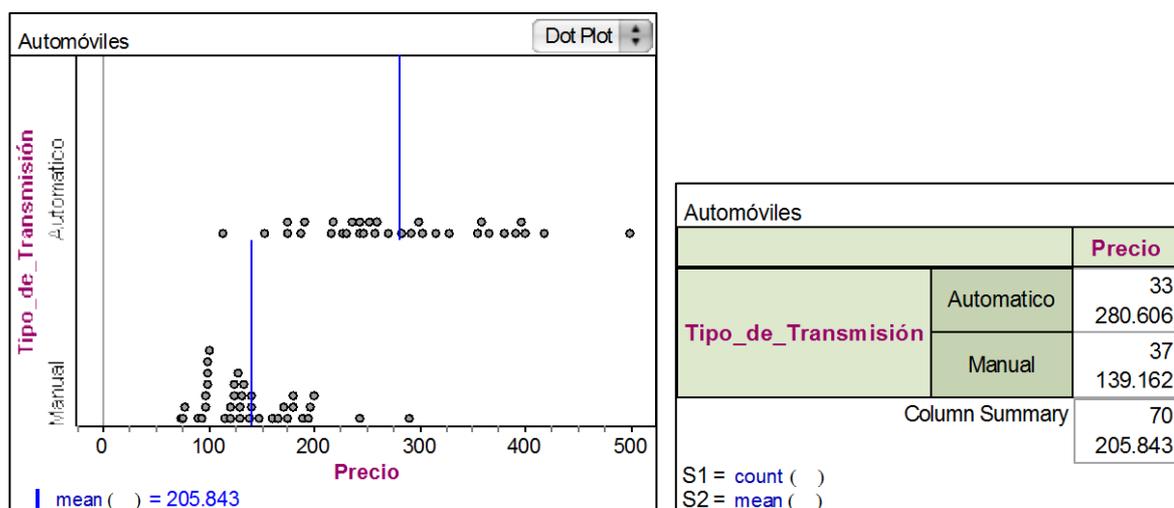


Figura 8 - Representaciones construidas por Bianca

Al representar ambas variables en una sola gráfica es posible apreciar que los automóviles automáticos son más costosos en promedio (280.60 mil pesos) que los manuales, cuyo promedio de costo es de 139.16 mil pesos.

Análisis de la Pregunta 4

La pregunta involucra dos variables cuantitativas. Las representaciones y cálculos realizados por los estudiantes se muestran en el cuadro siguiente:

Representaciones utilizadas	Cantidad
Diagrama de dispersión	9
Coefficiente de correlación	1
Diagrama de dispersión y coeficiente de correlación	22

Cuadro 7 - Representaciones utilizadas para el análisis de las variables *precio* y *potencia*

Del cuadro anterior se observa que los estudiantes, en su mayoría, se apoyaron en el diagrama de dispersión y el cálculo del coeficiente de correlación para responder la pregunta planteada. Aunque para efectos del cálculo de la correlación entre las variables, no es necesario distinguir entre variable de predicción y variable de respuesta, si lo es para el cálculo de la ecuación de regresión. De tal forma, en el uso del diagrama de dispersión 15 estudiantes cometieron errores al colocar incorrectamente en los ejes del diagrama las variables de predicción y de respuesta. Veamos el caso del estudiante Atanael, que coloca la potencia del automóvil en el eje vertical y el precio en el eje horizontal, cuando debió ser al revés, ya que el precio depende de la potencia.

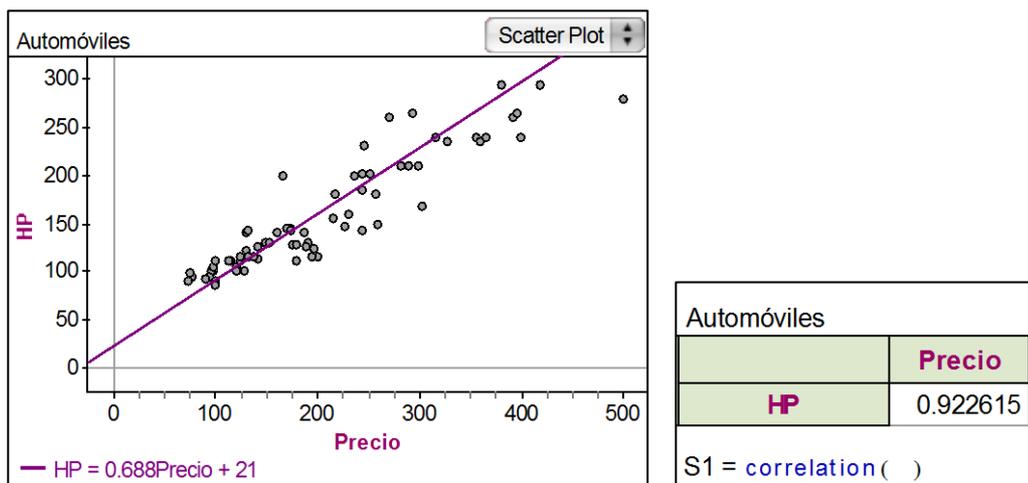


Figura 9 - Representaciones construidas por Atanael

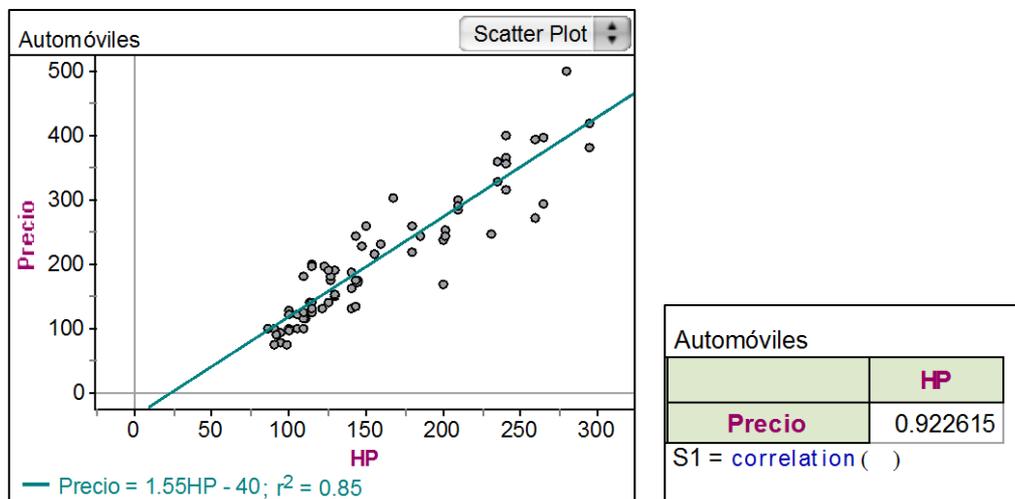


Figura 10 - Representaciones de datos construidas por Bianca

Al anteponer estas dos variables en una gráfica se observa que ésta es positiva con una intensidad fuerte además de que su coeficiente de correlación es de 0.922615 lo cual indica que se tiene un coeficiente muy fuerte ya que está muy cercano al 1 con lo cual se concluye que el precio del auto depende de la potencia (HP) de éste, es decir, entre más potencia tenga el auto mayor será su precio.

Manuel: La correlación es de 0.92, se aproxima mucho a 1 eso significa que la correlación es muy alta y de forma positiva eso indica que entre más HP "Potencia" tenga el automóvil mayor es el precio.

Joaquín: Al analizar las dos variables muestra que existe una correlación muy fuerte entre las dos variables con una correlación de 0.92, la cual indica que el precio depende de la potencia del automóvil, entre más potente sea el precio será más alto.

Se observa que en los estudiantes identifican correctamente el tipo de asociación que existe entre las variables y relacionan la intensidad con el valor del coeficiente de correlación. No obstante, se destaca el error que algunos estudiantes como Atanael cometieron en la gráfica, lo que trae como consecuencia que la ecuación de regresión obtenida sea incorrecta; sin embargo, el análisis no contempló la predicción de valores de la variable de respuesta.

Análisis de la Pregunta 5

Esta pregunta involucra una variable cualitativa y una cuantitativa. El propósito de la pregunta es ver si los estudiantes identifican a la marca como una variable que clasifica a los automóviles en grupos, ver si utilizan algún promedio por grupo (estructura global) o si se enfocan en valores individuales (valores mínimos) para decidir sobre las tres marcas más económicas. Las representaciones más sencillas que podían haber utilizado los estudiantes para responder la pregunta era una tabla descriptiva de medias o medianas por marca de vehículo y/o un diagrama de caja por marca. En cualquiera de los dos casos se podría decidir con facilidad las tres marcas más económicas. También se previó que utilizaran diagramas de puntos con medias marcadas sobre la gráfica para determinar visualmente el resultado. El análisis de las respuestas de los estudiantes nos muestra el siguiente cuadro con las representaciones utilizadas en el análisis:

Representaciones utilizadas	Cantidad
Diagramas de puntos	21
Tabla con medidas descriptivas	24
Diagramas de caja	2
Histogramas	4
Otras gráficas incorrectas	4
Tabla con medidas descriptivas y representación gráfica	22

Cuadro 8 - Representaciones utilizadas para responder la pregunta 5

La representación más utilizada por los estudiantes fue la tabla con medidas descriptivas; en algunos casos solo con conteos y medias, en otros casos con diversas medidas descriptivas como valores mínimos, máximos, cuartiles y desviación estándar. Es importante señalar que 22 estudiantes utilizaron estas tablas en complemento con una representación gráfica, y fue el principal elemento de decisión para responder la pregunta planteada. Por su parte, la representación gráfica más utilizada por fue el diagrama de puntos (marca contra precio) llegando en algunos casos a colocar líneas que indican el precio medio, lo que ayudó a los estudiantes a definir las tres marcas más económicas (ver figura 11) como se había previsto.

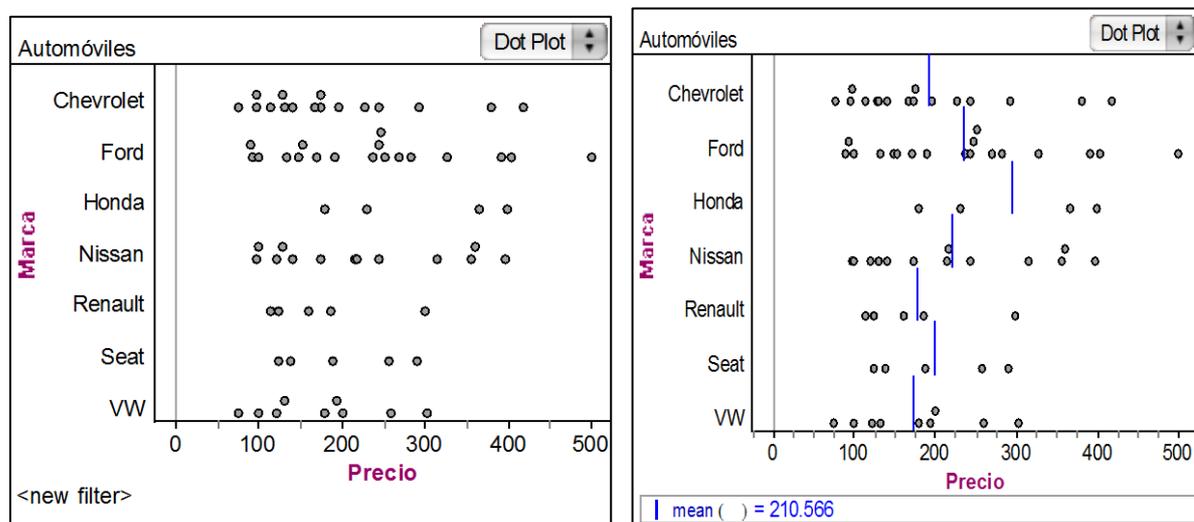


Figura 11 - Diagrama de puntos marca vs. Precio

De los 21 estudiantes que utilizaron el diagrama de puntos, 9 utilizaron la gráfica de lado izquierdo de la figura 11, por lo que tuvieron la necesidad de construir adicionalmente una tabla con medidas descriptivas entre las que se incluía la media y/o la mediana para poder identificar las tres marcas más económicas, ya que la gráfica por sí sola era insuficiente, 12 estudiantes utilizaron la gráfica del lado derecho y adicionalmente construyeron la tabla de medidas descriptivas aún cuando no era totalmente necesaria para responder la pregunta. A continuación se muestra la respuesta de Atanael:

La marca más barata sería VW con una media de 173.33, seguida de Renault con 176.2 y luego Chevrolet. Aún cuando esta última marca tiene el segundo auto más caro de todos, tiene muchos autos de precio bajo y su media de precios fue de 190.625

De lo anterior se observa que Atanael utiliza la tabla para definir el orden del precio de los automóviles y la gráfica para proporcionar información adicional al señalar que la marca Chevrolet, a pesar de tener el segundo vehículo más caro, se ubica como la tercera marca más

económica. Por su parte, otro estudiante (Daniel) utiliza un diagrama de puntos con líneas que identifican a la media y complementa su análisis con una tabla con valores medios (ver figura 12). Además de proporcionar las tres marcas más económicas, la gráfica le permite agregar otras características de los datos como es el caso de la variabilidad.

Las marcas más económicas son VW y RENAULT ya que sus autos más económicos van desde \$75,000 hasta \$113,000 y los más costosos apenas pasan los \$300,000, siendo que todas las otras marcas sobrepasan esas cantidades. El precio promedio de RENAULT es de \$176.2 y el de VW de \$173.3

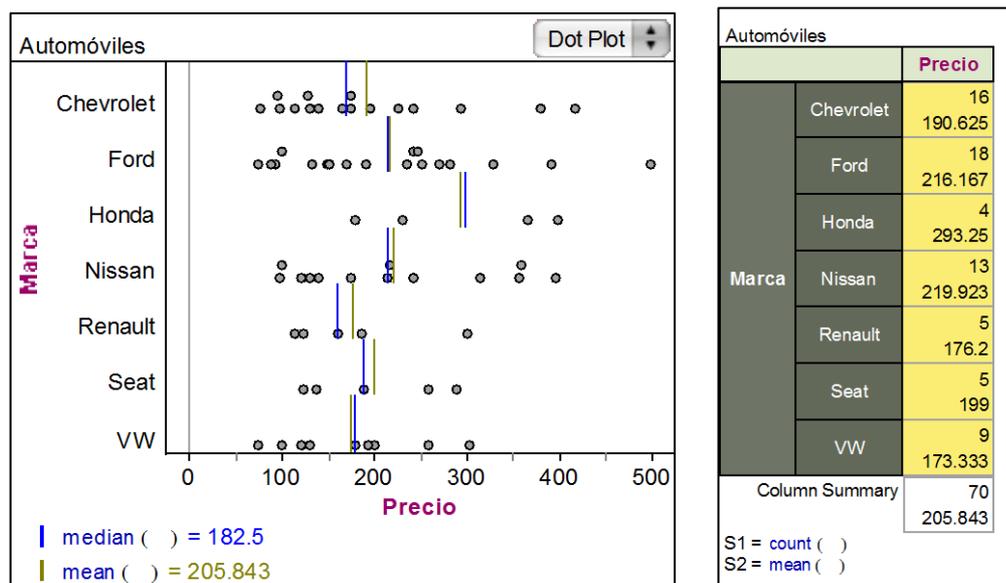


Figura 12 - Representaciones construidas por Daniel

El resto de los estudiantes que utilizaron estas representaciones, aún cuando identificaron correctamente el promedio como elemento adecuado para decidir sobre las marcas más económicas, no proporcionaron información adicional que las gráficas les proporcionaban, por lo que se deduce que la tabla con promedios fue el principal elemento de decisión, dejando de lado importantes propiedades como la variabilidad, los cuartiles, valores por encima o debajo de la media del precio. Por ejemplo, Nadia construyó una gráfica muy importante (ver figura 13), pero no la utilizó en su descripción, cuando le pudo haber servido para señalar que las marcas más económicas tenían la mayor parte de sus precios por debajo de la media global.

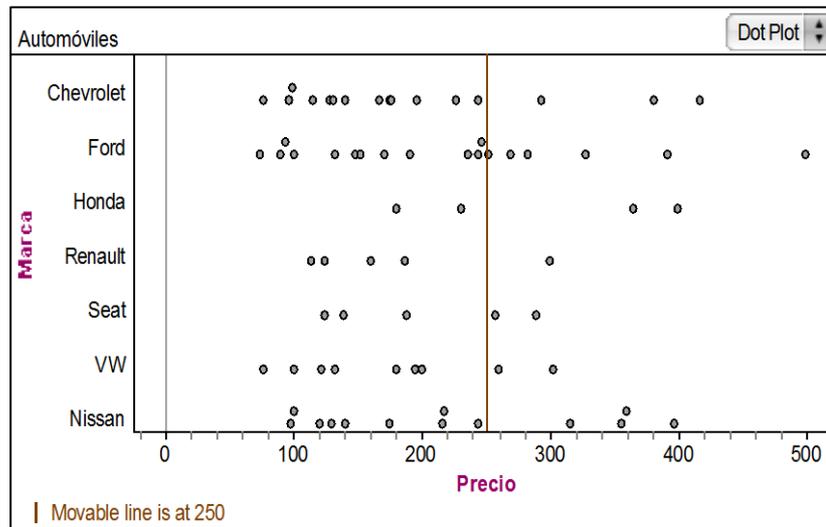


Figura 13 - Gráfica de puntos construida por Nadia

Otra representación, que fue utilizada en forma aislada y que en nuestro análisis *a priori* habíamos previsto que podría ser utilizada con mucha frecuencia, es el diagrama de cajas. Sin embargo, no ocurrió así y solo fue utilizada por dos estudiantes en complemento con una tabla descriptiva que no era totalmente necesaria. El diagrama de cajas será suficiente para decidir a través de la mediana las tres marcas más económicas, como se muestra la gráfica construida por Jesús (ver figura 14).

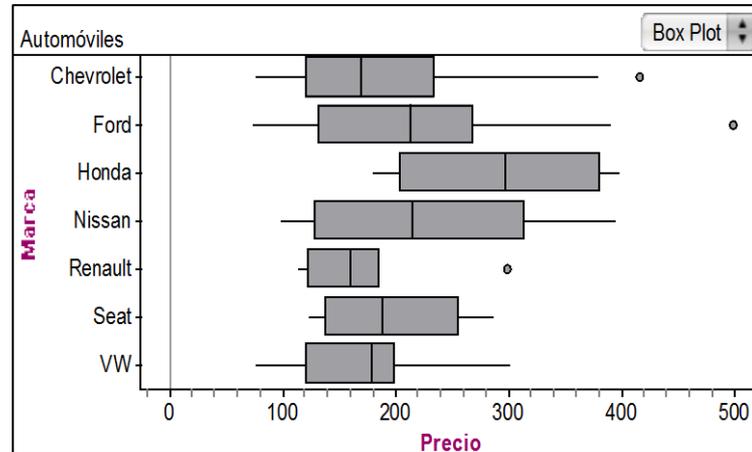


Figura 14 - Gráficas de caja construidas por Jesús

Finalmente, los histogramas también fueron utilizados, aunque estos no estaban previstos en el análisis *a priori*, pues por sí solos no proporcionan información útil para responder esta pregunta. Bianca construyó el siguiente histograma, aunque el análisis lo basó en la tabla no en el histograma.

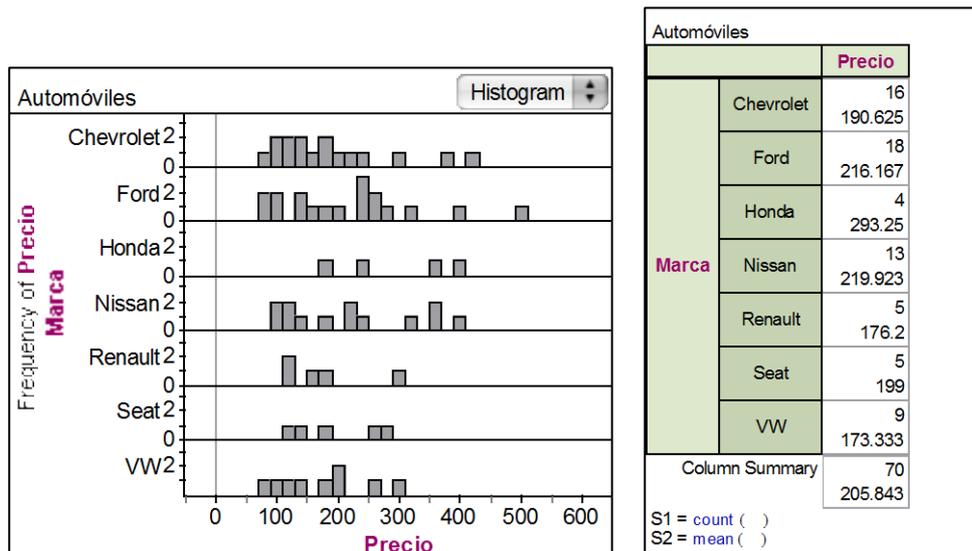


Figura 15 - Representaciones de datos construidas por Bianca

Me da como resultado que la más barata en promedio es la VW con 173.3 mil pesos, le sigue Renault con 176.2 mil pesos y por último Chevrolet con 190.625 mil pesos.

Para tener una visión global sobre las representaciones que utilizaron para responder las preguntas, se tiene el siguiente cuadro:

Representación posible de utilizar	Preguntas				
	1	2	3	4	5
Diagrama de barras	32				
Diagrama de cajas		3	7		2
Diagrama de puntos		11	15		21
Diagrama de barras acumuladas	1				
Histograma		9	4		4
Tabla resumen con medidas descriptivas	16	25	1	1	24
Diagrama de dispersión				9	
Alguna gráfica y tabla con medias descriptivas	15	24	22	22	22
Gráficas incorrectas		11	6	15	4

Cuadro 9 - Frecuencia con la que fueron utilizadas las representaciones previstas en el análisis a priori para cada pregunta

6 Conclusiones

El ambiente computacional en el cual los estudiantes realizaron el análisis de los datos facilitó el uso de diversas representaciones, sus transformaciones y el cálculo de medidas

descriptivas. Generalmente los estudiantes transitaban de una representación gráfica a otra para seleccionar la más adecuada al tipo de datos que estaban analizando. En la mayoría de los casos las gráficas fueron complementadas con tablas que contenían resultados de medidas descriptivas de los datos, aún cuando algunas veces una sola representación era suficiente para realizar el análisis de los datos; esto como consecuencia de la flexibilidad y el potencial amplificador del *software* utilizado.

Sin embargo, a pesar de la multiplicidad de representaciones que el *software* les proporcionaba para realizar un análisis detallado de los datos, se observó una tendencia a depender más de las tablas que de las gráficas, y en particular se hizo mayor énfasis en promedios (media o mediana), dada la familiaridad que tenían con estos conceptos de sus estudios previos. De tal forma, medidas descriptivas como los cuartiles, la variabilidad, forma y datos atípicos, fueron utilizados con menos frecuencia en el análisis de los datos; y si bien, en muchos casos fueron calculados por los estudiantes, con frecuencia solo fueron tomados de las tablas y descritos superficialmente y en forma aislada, sin llegar a realizar conexiones entre ellos para una interpretación más completa. Asumimos, por ello, que la interpretación de gráficas y tablas resultó una tarea compleja, dada la información diversa que proporciona cada representación.

Los resultados muestran que pasar de una visión local de los datos a una comprensión global no fue tarea sencilla para los estudiantes, a pesar de contar con una herramienta cognitiva (*software*) con un poder amplificador y reorganizador que les permitía el uso de múltiples representaciones y el paso flexible entre ellas, para extraer la información de los datos. Sin embargo, en la última actividad, aún cuando sus descripciones incorporan más propiedades de los datos, consideramos que la comprensión es aún insuficiente, en tanto no se establecen relaciones de mayor nivel cognitivo entre dichas propiedades.

Dentro de las implicaciones del presente estudio podemos señalar que un análisis adecuado de un conjunto de datos que permita comprender su comportamiento y extraer toda la información estadística – contrario a lo que muchos profesores piensan – es una tarea compleja que requiere poner en juego un conocimiento profundo de los conceptos estadísticos. En este proceso la tecnología juega un papel importante, y particularmente si ésta reúne elementos de una herramienta cognitiva, que permita a los estudiantes explorar variables de una forma flexible a través de sus representaciones.

Referencias

- BAKKER, A.; DERRY, J.; KONOLD, C. Technology to support diagrammatic reasoning about center and variation. In: 7TH INTERNATIONAL CONFERENCE ON TEACHING STATISTICS. Salvador Bahía. Brazil. Voorburg: International Statistical Institute, 2006. CD-ROM.
- BEN-ZVI, D. Seventh grade students' sense making of data and data representations. In: 6TH INTERNATIONAL CONFERENCE ON TEACHING STATISTICS, Cape Town South Africa. Voorburg: International Statistical Institute, 2002. CD-ROM.
- BEN-ZVI, D.; ARCAVI, A. Junior high school students' construction of global views of data and data representations. **Educational Studies in Mathematics**, Dordrecht, v. 45, n. 1, p. 35-65. March. 2001.
- BIHLER, R. Students' difficulties in practicing computer-supported data analysis: Some hypothetical generalizations from results of two exploratory studies. In: GARFIELD, J.; BURRILL, G. (Eds.). **Research on the role of technology in teaching and learning statistics**. Voorburg. The Netherlands: International Statistics Institute, 1997, p. 169-190.
- BIGGS, J.; COLLIS, K. **Evaluating the Quality of Learning: the SOLO taxonomy**. New York: Academic Press. 1982.
- BURRILL, G.; FRANKLIN, CH.; GODBOLD, L.; YOUNG, L. **Navigating through Data Analysis Analysis in Grades 9-12**. Reston: VA. National Council of Teachers of Mathematics. 2003.
- COBB, P.; MCCLAIN, K. Principles of Instructional Design for Supporting the Development of Students' Statistical Reasoning. In: BEN-ZVI, D.; GARFIELD, J. (Eds.). **The Challenge of Developing Statistical Literacy, Reasoning, and Thinking**. Dordrecht, The Netherlands: Kluwer Academic Publishers. 2004. p. 375-396.
- FINZER, W. Fathom: Dynamic data exploration. [Computer Software]. Emeryville, CA: Key Curriculum Press. 2002.
- FRANKLIN, C.; KADER, G.; MEWBORN, D.; MORENO, J.; PECK, R.; PERRY, M.; SCHEAFFER, R. **A Curriculum Framework for K-12 Statistics Education**. GAISE report. 2005. Disponible in: <www.amstat.org/education/gaise/>. Access in: 03 Mai. 2013.
- GARFIELD, J. The Challenge of Developing Statistical Reasoning" **Journal of Statistics Education**. Alexandria, VA, v. 10, n. 3, 2002. Disponible en: www.amstat.org/publications/jse/v10n3/garfield.html. Acceso en: 10 Nov. 12
- GARFIELD, J.; BEN-ZVI, D. **Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice**. Springer. 2008.
- INZUNSA, S.; VIZCARRA, F. Un estudio sobre la caracterización del razonamiento estadístico de estudiantes de preparatoria: el caso de los promedios y las gráficas. In: MEMORIAS DEL CONGRESO NACIONAL DE ENSEÑANZA DE LAS MATEMÁTICAS, 22th, 2009. **Anales...** Tuxtla Gutiérrez México. Tuxtla Gutiérrez México: Asociación Nacional de Profesores de Matemáticas (ANPM). 2009.
- INZUNSA, S.; JUAREZ, J. A. High School Teachers' Reasoning about Data Analysis in a Dynamic Statistical Environment. In: INTERNATIONAL CONFERENCE ON TEACHING STATISTICS, 8th, 2010. **Anales...** Ljubljana, Slovenia. Voorburg: International Statistical Institute. 2010. CD-ROM.
- KONOLD, C.; MILLER, C. **TinkerPlots: Dynamic data exploration. [Computer software]** Emeryville, CA: Key Curriculum Press. 2005.



LAJOIE, S.; ROMBERG, T. A. Identifying an Agenda for Statistics Instruction and Assessment in Grades K-12. In: LAJOIE, S. P. (Ed.). **Reflections on Statistics: Learning, Teaching, and Assessment in Grades K-12**. Mahwah New Jersey, USA: Lawrence Erlbaum Associates. 1998. p. vii-xxi.

MINISTRY OF EDUCATION. **The New Zealand Curriculum**. Wellington, New Zealand: Learning Media Limited. Ministry of Education. 2007.

MINISTERIO DE EDUCACION Y CIENCIA (MEC). **Real Decreto 1631/2006, de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria**. Madrid. BOE núm. 5 de 05/01/2007. 2006.

MONTEIRO, C.; ASSEKER, A.; CARVALHO, L.; CAMPOS, T. Student teachers developing their knowledge about data handling using TinkerPlots. In: 8TH INTERNATIONAL CONFERENCE ON TEACHING STATISTICS. Ljubljana, Slovenia. **Proceedings...** Voorburg: International Statistical Institute. 2010. CD-ROM.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS (NCTM). **Principles and Standards for School Mathematics**. Reston VA. 2000.

PEA, R. Cognitive Technologies for Mathematics Education. In SCHOENFELD, A. (Ed.). **Cognitive Science and Mathematics Education**. Hillsdale New Jersey, USA: Lawrence Erlbaum Associates Publishers. 1987. p. 89-122.

REFORMA (2010). Disponible en: <www.reforma.com.mx>. Acceso en: 10 Nov. 12

REVISTA DEL CONSUMIDOR (PROFECO). Disponible en: <<http://revistadelconsumidor.gob.mx/>>. Acceso en: 10 Nov. 12

TUKEY, J. W. **Exploratory Data Analysis**. Addison-Wesley. 1977.

VIZCARRA, F.; INZUNSA, S. Desarrollo del razonamiento estadístico con TinkerPlots. **Memorias de la Conferencia Interamericana de Educación Matemática**. 13th. Recife Brasil. http://www.gente.eti.br/lematec/CDS/XIIICIAEM/?info_type=home&lang_user=es. 2011.

WILD, C. J.; PFANNKUCH, M. Statistical Thinking in Empirical Enquiry. **International Statistical Review**, The Netherlands, v. 67, n. 33, p. 223-265. 1999.

**Submetido em Junho de 2013.
Aprovado em Fevereiro de 2014.**