

2

Elaboración de estadísticas de vacantes publicadas en internet

Una experiencia en Ecuador

Diego Benítez⁺

Sebastián Lucero^{*}

Ana M. Pazmiño^{*}

RESUMEN

La presente Nota Metodológica documenta un ejercicio de elaboración de estadísticas de vacantes, a partir de los anuncios de empleos que se publican en el Internet, usando una herramienta de web scraping. Se describen los aspectos más relevantes del proceso; y se muestra, a manera de ejemplo, los principales resultados correspondientes a los meses de septiembre y octubre de 2016.

Palabras clave: webscraping, internet, vacantes

Descargo de responsabilidad: Las opiniones e interpretaciones expresadas en este documento pertenecen a los autores y no reflejan el punto de vista oficial del Instituto Nacional de Estadística y Censos (INEC). El INEC no garantiza la exactitud de los datos que figuran en el documento.

+ Consultores del Banco Mundial. Los autores agradecen el apoyo financiero del *Jobs Umbrella Trust Fund* para la realización de este documento así como también para muchas otras actividades orientadas al fortalecimiento de la producción de estadísticas laborales en Ecuador. Lo expresado en este documento no representa la posición del INEC, del gobierno del Ecuador, ni del Grupo Banco Mundial. Cualquier error es de los autores. Correspondencia a través de dieben79@yahoo.com. Se agradece también la orientación proporcionada por Jeisson Cárdenas para la realización del ejercicio. Con la finalidad de contribuir al debate técnico respecto al mercado laboral, orientado a la formulación de mejores políticas públicas, el Instituto Nacional de Estadísticas y Censos (INEC) y el Banco Mundial (BM) emprendieron un proyecto conjunto denominado "Data, Diagnostics and Institutional Innovations for Jobs in Ecuador" en el que uno de sus componentes es el fortalecimiento de las estadísticas relacionadas con la temática. En este contexto, se decidió explorar la posibilidad de elaborar estadísticas de vacantes a partir de anuncios de empleos que se publican en internet, a manera complementaria de otras estadísticas que cuenta o que está desarrollando el INEC.

1. Introducción

Actualmente, millones de gigabytes de información se suben al internet diariamente. Muchos profesionales como periodistas, investigadores, analistas de datos, agentes de ventas, desarrolladores de software, entre otros, acuden a esta información, y con técnicas de “*copy and paste*” la organizan en hojas de cálculo y forman parte de sus informes y presentaciones. Sin embargo, cuando ésta es muy grande, tal es el caso si la información a consultar tiene fines estadísticos, este trabajo manual puede volverse demasiado tedioso.

Para superar este inconveniente, se desarrollaron técnicas de *web scraping* que permiten extraer información de los sitios web de forma eficiente y automática, y convertirla en formatos más estructurados y fáciles de usar (Castrillo-Fernández, 2015). Los usos del *web scraping* van desde el análisis de comportamiento hasta aplicaciones de economía política, e incluso índices de precios complementarios. Una de sus aplicaciones más destacadas es el estudio de la demanda laboral, como es el caso de Australia, donde su oficina nacional de estadística cuenta con el Índice de Vacantes de Internet (*Internet Vacancy Index*) (Reimsbasch-Kounatze, 2015), el cual muestra mensualmente la variación de vacantes publicadas en las principales bolsas de empleo *online*. A nivel regional, en Colombia, Cárdenas et al. (2015) realizaron una metodología para el análisis de demanda laboral mediante datos de Internet que, pese aún no estar institucionalizada, sentó un precedente y ha sido un elemento clave en el desarrollo del presente trabajo.

Inspirados en la experiencia colombiana, el INEC, con el apoyo del Banco Mundial, impulsó la evaluación de la posibilidad de elaborar estadísticas de vacantes como parte del sistema de información relacionado con el mercado laboral en Ecuador. El presente artículo tiene por objetivo documentar esta experiencia. Para lograr este objetivo, a manera de justificación, primero se explica qué parte de la información acerca de la demanda laboral puede cubrirse con los anuncios de empleo que se publican en internet. Segundo, se hace una breve revisión de otras experiencias en el uso del *web scraping* con fines estadísticos. Tercero, se describe el proceso técnico de extracción de la información en el caso ecuatoriano. Cuarto, se presenta los principales resultados correspondientes a los meses de septiembre y octubre de 2016. Finalmente se exponen las principales conclusiones y recomendaciones.

2. Justificación

Si bien es deseable una situación en la que tanto la oferta como la demanda en el mercado de trabajo se encuentren en un punto de equilibrio óptimo, es decir, dónde todos los empleadores tienen cubiertas sus plazas de trabajo y las personas que desean tener un empleo así lo hacen, esto no siempre sucede así y mucho menos en tiempos dónde la actividad económica en general tiende a disminuir. Uno de estos desequilibrios es el desempleo, mismo que puede ser originado por factores estructurales, friccionales o cíclicos.

Estructuralmente, el desempleo se origina por varios factores, uno de ellos es las brechas entre las habilidades que requieren las empresas y las que pueden ofrecer los trabajadores en un momento dado. Generalmente, los saltos tecnológicos hacen que las habilidades de algunos trabajadores queden obsoletas. Las medidas para corregir el desempleo estructural tardan un largo plazo en hacer efecto. Por su parte, los factores friccionales del desempleo corresponden al tiempo que tarda un trabajador en transitar de un empleo a otro; este tipo de desempleo podría reducirse tanto si empleadores como trabajadores contaran con información adecuada. Finalmente, los factores cíclicos están relacionados con la actividad económica general, y el desempleo ocurre cuando la demanda agregada se contrae a tal punto en que algunas empresas precinden de sus trabajadores.

La intermediación laboral ayuda a resolver las brechas y distorsiones propias del desempleo friccional; entendida como tal a las instituciones y procesos que facilitan el intercambio de información entre empleadores y trabajadores con el fin de reducir el tiempo que una vacante toma en ser llenada. La proliferación del uso de tecnologías de información y comunicación han permitido la aparición de nuevos esquemas de intermediación laboral más ágiles y versátiles, como es el caso de las bolsas de empleo *online*. Kuhn (2011), provee evidencia sobre el importante rol de las herramientas *online* para emparejar trabajadores y plazas de trabajo.

Por lo expuesto, el estudio de la información que se publica en las bolsas de empleo puede proveer elementos útiles para entender una parte importante del mercado de trabajo. Esto es importante porque potencialmente permite llenar un vacío de información respecto al mercado laboral, que las fuentes tradicionales, enfocadas principalmente a la oferta laboral, no cubren.

Específicamente, en Ecuador la principal fuente de información sobre el mercado laboral es la Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU), a través de la cual se calculan los principales indicadores laborales vinculados con la condición de actividad como las tasas empleo adecuado, inadecuado y desempleo, entre otros (INEC, 2016). Por el lado de la demanda de trabajo pueden citarse fuentes como los Estudios Mensuales de Opinión Empresarial (EMOE) del Banco Central del Ecuador (BCE), o las Encuestas Industriales y el Laboratorio de Dinámica Laboral y Empresarial (LDLE)¹ del INEC (Benítez, Espinoza, Grijalva, Rivadeneira y Oviedo (2016).²

La información de los anuncios de los empleos que se publican en internet permitiría estudiar una parte del mercado laboral que aún no ha sido explorada por las otras fuentes. Con información sistematizada sobre las ofertas de plazas de trabajo se podría conocer el perfil de profesionales que las empresas requieren, incluyendo detalles como: habilidades, localización de la vacante, salario, experiencia mínima, tipo de contrato, entre otros. Al utilizar como técnica el *web scraping*, la información que se obtiene es oportuna, y su costo de levantamiento es relativamente bajo. En este sentido, Carnevale, et al. (2014) señalan como ventajas el bajo costo para su recolección, el permitir contar con el detalle de ocupaciones, industria y ubicación, el mostrar la demanda de habilidades, el posibilitar realizar monitoreos y notar tendencias, y el ser útil para mejorar la política educativa.

Indudablemente, existen también ciertos reparos a la obtención de estadísticas por esta vía. Muchos de estos problemas se relacionan con la calidad, confiabilidad y representatividad de la información obtenida (Stefanik, 2012). Estas preocupaciones son especialmente pertinentes, dado que las principales herramientas cuantitativas del análisis económico están basadas en el uso de la estadística inferencial. Sin embargo, en un horizonte lejano, hay que considerar que en la medida en que la población va digitalizándose por completo, las técnicas de muestreo pueden incluso llegar a ser obsoletas (Askitas & Zimmermann, 2015).

1 El LDLE es un sistema de registros administrativos con fines estadísticos. Esto implica que se nutre de cierta información de empresas que reportan información a administraciones públicas como la oficina de impuestos o el instituto de seguridad social, SRI o IESS en el caso particular del Ecuador.

2 Otras fuentes se están desarrollando en el marco de la cooperación INEC-BM, como son los casos del módulo de demanda laboral de corto plazo anclado al Sistema de Indicadores de la Producción (SIPRO), o el módulo de demanda laboral estructural anclado a las encuestas empresariales; pero ambas se encuentran aún en fase de desarrollo.

3. Otros usos de web scraping con fines académicos

Pese a los problemas que presenta, el *web scraping* se viene usando de manera creciente en diferentes disciplinas. En este sentido, la Economía no constituye una excepción. Según Edelman (2012) la toma sistemática de información de internet ha sido utilizada en varios estudios acerca de: historia del pensamiento económico (Azar, 2007), microeconomía (Bajari & Hortacsu, 2003), finanzas (Antweiler & Murray, 2004), desarrollo económico (Seamans & Zhu, 2010), entre otros.

Azar (2007), examina las tendencias en los tiempos de respuesta de las revistas científicas a los borradores de artículos que reciben, recogiendo información de tiempos de respuesta de diferentes páginas web. Bajari & Hortacsu (2003), toman información de las ventas de monedas coleccionables en eBay para analizar el comportamiento de los postores en las subastas. Antweiler & Murray (2004), analizan las discusiones de los foros *online* de los mercados de valores y encuentran que las discusiones ayudan a predecir precios. Finalmente, Seamans & Zhu (2010) exploran la forma en que los periódicos locales adaptan sus tarifas de suscripción, de publicidad y de anuncios clasificados frente a la entrada de páginas como *Craigslist*³.

Uno de los usos más destacados de la construcción de estadísticas a partir de Internet tiene que ver con la estimación de índices de precios. En la medida en que cada vez más productos son vendidos *online*, el Internet se convirtió en una fuente rica para obtener información de precios. Esto lo supo ver bien Alberto Cavallo que en 2007 presentó su primer desarrollo utilizando precios en línea de Argentina, Chile, Brazil, y Colombia (Cavallo, 2009). Sus exitosos resultados le permitieron expandir su colección de información a más de 50 países, dando inicio a la iniciativa académica Billion Price Project (BPP) que actualmente es manejada por PriceStats. En 2012, finalmente logró demostrar que las cifras oficiales de Argentina estaban siendo manipuladas por lo que su sitio *online* www.inflacionverdadera.com pasó a ser la fuente más creíble de inflación (Cavallo, 2012).

3 *Craigslist* es un sitio web de anuncios clasificados con secciones dedicadas al empleo, vivienda, contactos personales, ventas, ítems, servicios, comunidad, conciertos, hojas de vida, y foros de discusión.

La utilización de información tomada a partir de Internet en la investigación y análisis del mercado laboral difiere según los objetivos y enfoques perseguidos. Reimsbasch-Kounatze (2015) expone un par de casos significativos: por ejemplo, en Estados Unidos, la actividad de creación de plazas de trabajo es monitoreada a nivel nacional, regional, estatal y metropolitano a través de los anuncios de trabajo condensados en la forma de un índice compuesto denominado *Help-Wanted Index*.⁴

Otro caso importante está en Australia, en donde se da seguimiento a la variación mensual del total de anuncios de las tres bolsas de empleo más grandes a través del *Internet Vacancy Index* (IVI). El IVI incluye información de al menos 350 tipos de ocupaciones para todas sus 37 regiones. En la Unión Europea (UE), una fuente potencial para información de vacantes online es el sitio web EURES (European Employment Services), el cual almacena las vacantes laborales de todos los países de la UE en una plataforma estandarizada (Kurekova, et al., 2015). A nivel regional, Colombia es pionero en el uso de vacantes *online* para estimar la demanda laboral. Cárdenas, et al. (2015) presentan una metodología para analizar la demanda laboral mediante la extracción de información *online* de plazas de trabajo vacantes.

4. La experiencia en Ecuador

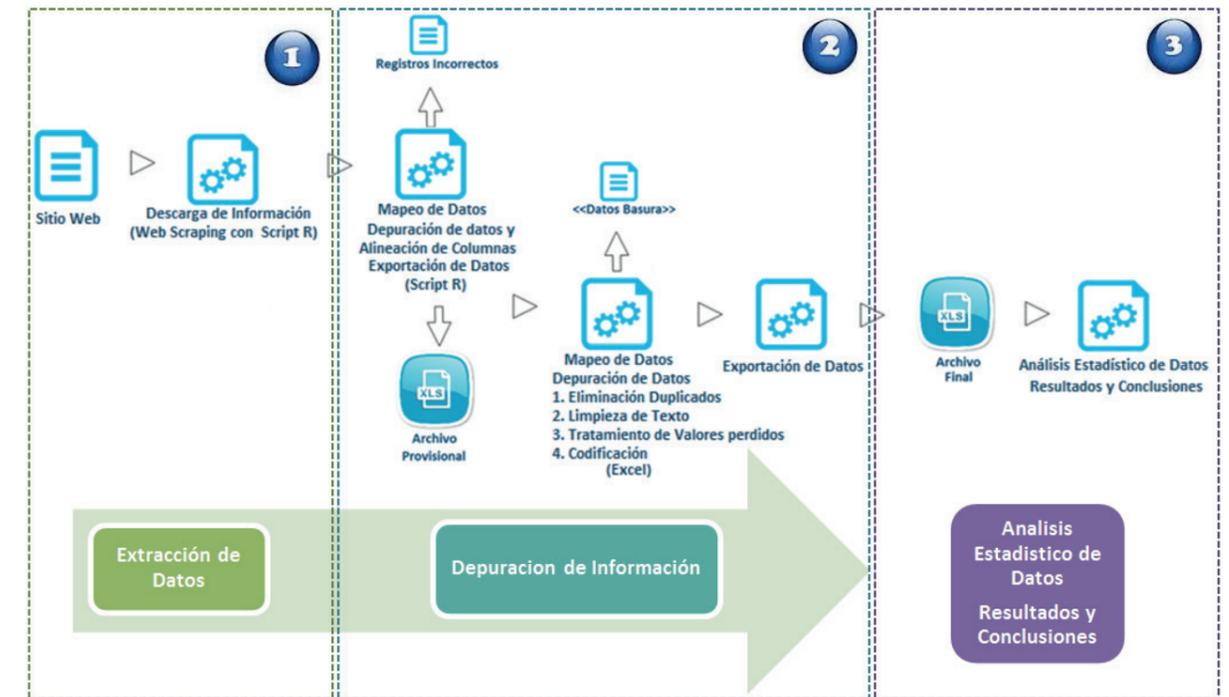
Específicamente, la experiencia que aquí se presenta consistió en elaborar estadísticas de las vacantes que se publican en internet. En Ecuador existen distintas bolsas de empleo online entre las que destacan una pública (Red Socioempleo⁵), y algunas privadas como: CompuTrabajo⁶, Multitrabajos⁷, Porfinempleo⁸, entre otras. De ellas, para este ejercicio se eligió CompuTrabajo, por ser considerada una de las más utilizadas en el medio. En Ecuador, esta bolsa publica mensualmente más de 2 000 anuncios que incluyen campos de información como el nombre de la empresa, nombre del cargo, tipo de contrato, salario, educación mínima, entre otros elementos.

Es probable que un estudio que pretenda abarcar el universo de los anuncios de empleo, requiera incorporar varias de las bolsas disponibles. Sin embargo, resulta difícil armar una base de datos con el total de anuncios, dado que, no se conoce si los anuncios publicados en las distintas páginas web se complementan o se superponen, y en qué medida lo hacen, lo cual constituye un impedimento para combinar la información de dos o más bolsas distintas.

Una solución a este problema podría ser identificar si dos o más anuncios que se han publicado en distintos medios corresponden realmente a la misma vacante. Esto podría hacerse mediante la comparación de los campos de los anuncios, siempre que puedan establecerse patrones que permitan determinar la similitud y sugerir la superposición del dato. Sin embargo, esto no necesariamente es fácil de implementar, para conocer su factibilidad se requiere de un desarrollo tecnológico adicional. Una solución complementaria o alternativa consiste en utilizar encuestas a las empresas para conocer el contexto en el cual ellas difunden sus vacantes disponibles, con lo cual podría saberse, por ejemplo, el porcentaje de empresas que usan un medio u otro. Esta segunda opción debe considerar el costo inherente a un levantamiento de campo, por lo que, para ahorrar costos, podría aprovecharse alguna encuesta en marcha para incluir un conjunto de preguntas en este sentido.

Bajo estas consideraciones, la experiencia que aquí se reporta se centró en elaborar estadísticas de plazas de trabajo publicadas en la página web de CompuTrabajo. El proceso consiste en: i) utilizar herramientas de *web scraping* para descargar periódicamente la información de anuncios de empleo que aquí se publican, ii) depurar esta información para almacenarla en una base de datos estructurada, y finalmente, iii) obtener estadísticas a partir de la misma. El gráfico 1 describe este proceso.

Gráfico 1. Proceso de construcción de la base de información



4.1. Descarga de la información

En el proceso de captura y almacenamiento de la información, *web scraping*, se utilizan algoritmos de interpretación que permiten modelar la estructura de información de las páginas web para poder transformar datos sin estructura en bases de datos estructuradas. Este proceso se puede explicar más fácilmente de la siguiente manera: piénsese en una página web como un conjunto de campos de información ordenados donde cada campo de información puede ser identificado en el código fuente de la página mediante la interpretación del lenguaje HTML, técnicamente, este procedimiento se lo denomina análisis del esquema HTML. En este sentido, conociendo como se presenta la información en la web se puede extraer información de cada campo de información. Este proceso se programa en algoritmos para que sea realizado de manera repetitiva y estandarizada. Sin embargo, se debe tener presente que cualquier cambio en la forma de presentar la información, modifica el esquema HTML de la página y perjudica los códigos de programación *web scraping* y, en consecuencia, se requiere un esfuerzo de supervisión permanente.

La información extraída, es almacenada en una matriz donde las filas corresponden a los anuncios y las columnas a sus características asociadas (campos de información). Dado que los campos de información de CompuTrabajo son semiestructurados, es decir, no guardan siempre el mismo ordenamiento de la información para todos los anuncios, se identificó, en la base descargada, la existencia de atributos que no corresponden con la variable que debería. Para alinear correctamente todos los campos, se cuenta con un algoritmo capaz de corregir estos problemas, sin embargo, algunas observaciones por sus particularidades deben ser alineadas manualmente. Finalmente, posterior a la etapa de depuración, la base de datos se exporta a una hoja de cálculo.

La propuesta busca generar información mensual de vacantes. Para esto, se estableció que la extracción de información se realice cada 10 días con el fin de garantizar que todos los anuncios publicados en el mes de referencia han sido capturados adecuadamente. (Al momento, se cuenta con bases de datos para los meses de septiembre y octubre 2016).

⁴ El índice es elaborado por el *Conference Board*, una organización no gubernamental y *think tank* privado.

⁵ <http://www.socioempleo.gob.ec/>

⁶ <http://www.computrabajo.com.ec/>

⁷ <http://www.multitrabajos.com/>

⁸ <http://www.porfinempleo.com/>

4.2. Depuración de la data

En la depuración de la información extraída, se siguen una serie de pasos fundamentales previo a la obtención de la base de datos a ser analizada. Estos pasos tienen que ver con: 1) Eliminación de registros duplicados, 2) Limpieza de caracteres basura, 3) Tratamiento de valores *missing*, 4) Codificación de variables. A continuación, se describe cada uno de estos pasos:

- 1. Eliminación de observaciones duplicadas:** El primer paso en el proceso de limpieza es la eliminación de valores duplicados. Dado que la recolección y almacenamiento de información se realiza cada diez días, al momento de consolidar en una única base se detectó registros repetidos. Para solucionar esto, sobre la base consolidada se utiliza la función 'Quitar duplicados' de Microsoft Excel, la cual suprime todas las observaciones que sean duplicadas.
- 2. Limpieza del texto:** Una vez que se cuenta con registros únicos, se procede a realizar la limpieza de texto en todos los campos de información, esto es, remover todos los caracteres que no aportan o no guardan coherencia con la descripción de la variable. Por tratarse de extracción de información en lenguaje HTML, al momento de la captura de datos existen <<datos basura>>. Algunos de estos caracteres se refieren a signos de puntuación,

conjunciones y palabras no relacionadas. Este procedimiento se lo realiza con técnicas de minería de texto.

- 3. Tratamiento de valores perdidos:** El criterio utilizado para llenar los campos vacíos de información fue estrictamente basado en la observación y análisis de la variable FUNCIONES (tareas y competencias requeridas para ocupar la vacante) que, contiene generalmente una amplia descripción de la vacante y permite detectar palabras clave para completar los vacíos de información presentes en el resto de variables.
- 4. Codificación:** Para el análisis de la información y facilitar la comparabilidad es necesario codificar ciertos campos. Estos campos se refieren en específico a las variables capturadas del cargo y la localidad. Respecto al cargo, la codificación se hizo mediante el contraste con las ocupaciones registradas a 8 dígitos en la Clasificación Internacional Uniforme de Ocupaciones (CIUO-08). En cuanto a la localidad (que reporta el texto de provincia y ciudad), se separó su texto en "PROVINCIA" y "CIUDAD"; luego, empatando con la División Político Administrativa (DPA) se obtienen los códigos de provincia y cantón.

Finalmente, concluido todo el proceso de depuración, se logra contar con una base de datos lista para ser analizada, obtener resultados y emitir conclusiones. A continuación, se muestra la descripción de la base para cada una de sus variables (Tabla 1):

Tabla 1. Estructura de la base de datos

Variable	Tipo	Descripción
EMPRESA	Alfanumérico	Nombre de la empresa
DESC. OFERTA	Alfanumérico	Nombre de la vacante tal como se reporta
AREA	Alfanumérico	Nombre de la vacante tal como se reporta
CARGO	Alfanumérico	Nombre de la vacante
CODIGO CIUO	Código numérico	Código CIUO correspondiente con el cargo
DESCRIPCION CIUO	Alfanumérico	Nombre de la ocupación según el CIUO-08 a 8 dígitos
JORNADA	Categoría	Jornada de trabajo (tiempo completo, parcial, por horas, otros)
CONTRATO	Categoría	Tipo de contrato (indefinido, determinado, por obra, otros)
SALARIO	Alfanumérico	Salario
DESC EMPRESA	Área de texto	Descripción de la actividad de la empresa
LOCALIDAD	Alfabética	Ubicación de la vacante por cantón-provincia

DPA-PRV	Código numérico	Código de la provincia conforme DPA
PROVINCIA	Alfabética	Nombre de la provincia donde se localiza la vacante
DPA-CNT	Código numérico	Código de la ciudad conforme DPA
CIUDAD	Alfabética	Nombre de la ciudad donde se localiza la vacante
FUNCIONES	Área de texto	Tareas y competencias requeridas para ocupar la vacante
FECHA CONTRATACION	Fecha	Fecha de posible vinculación
FECHA PUBLICACION	Fecha	Fecha de publicación de la vacante
VACANTES	Numérica	Número de puestos de trabajo por vacante
EDUCACION MINIMA	Categoría	Nivel de estudios mínimo para acceder al cargo (bachillerato, universidad, postgrado, otros)
EXPERIENCIA	Numérica	Experiencia en años requerida
DIS. VIAJE	Dicotómico	Disponibilidad para viajar
IDIOMA	Alfabética	Idiomas requeridos para aplicar
SEXO	Categoría	Preferencia de sexo para la vacante (hombre, mujer, ambos sexos, no específica)

Fuente: CompuTrabajo, elaboración de los autores.

5. Resultados

A continuación, se presenta algunas estadísticas obtenidas del ejercicio. El análisis consiste en analizar: distribuciones para el total de observaciones, variaciones entre los meses de septiembre y octubre, y algunas relaciones importantes entre variables.

Del proceso de descarga y depuración se obtiene una base de datos que, entre septiembre y octubre de 2016, consta de 4.385 registros (anuncios) correspondientes a 13.271 vacantes (plazas de trabajo vacías). Esta diferenciación entre anuncios y vacantes se hace porque ocurre con cierta frecuencia que las empresas requieren contratar más de una vacante en un mismo anuncio, algo que puede verse más fácilmente, por ejemplo, en una empresa de limpieza que requiere diez personas para el cargo de asistente de limpieza. La Tabla 2 muestra el número de anuncios y de vacantes que se publicaron en CompuTrabajo en los meses de septiembre y octubre.

Tabla 2. Anuncios y vacantes

	2016		
	Septiembre	Octubre	Total
Anuncios	2.227	2.158	4.385
Vacantes	6.491	6.780	13.271

Fuente: CompuTrabajo, elaboración de los autores.

Si bien en octubre hay menos anuncios, existe un aumento en el número de vacantes, debido a que en octubre se tiene más anuncios que ofertan 10 o más plazas de trabajo (y menos anuncios con una sola vacante). La mayoría de anuncios, cerca de un 60%, ofrece una única plaza de trabajo, mientras que, en el otro extremo, hay un 7% de anuncios que presentan 10 o más plazas de trabajo vacantes. La Tabla 3 muestra la frecuencia de los números de vacantes por anuncio en los meses de septiembre y octubre.

Tabla 3. vacantes por anuncio

Número de puestos por anuncio	Sep	Oct	Total	%
1	1.345	1.299	2.644	60,3
2	377	373	750	17,1
3	150	140	290	6,6
4	58	63	121	2,8
5	111	91	202	4,6
6	23	21	44	1,0
7	5	2	7	0,2
8	9	9	18	0,4
9	0	2	2	0,0
10+	149	158	307	7,0
Total	2.227	2.158	4.385	100

Fuente: CompuTrabajo, elaboración de los autores.

A nivel geográfico, en la Tabla 4 se puede observar que la gran mayoría de vacantes se concentra en las tres principales ciudades del país. De esta manera, Quito aglutina un 51% de vacantes, seguido por Guayaquil con un 31% y Cuenca con apenas un 3%. En ambos meses no se aprecia ningún cambio significativo en la clasificación de ciudades.

Tabla 4. Vacantes por ciudad

Ciudad	Sep	Oct	Total	%
Quito	3.389	3.313	6.702	50,7
Guayaquil	1.973	2.113	4.086	30,9
Cuenca	179	203	382	2,9
Ambato	156	85	241	1,8
Machala	83	107	190	1,4
Santo Domingo	79	92	171	1,3
Rumiñahui	60	68	128	1,0
Portoviejo	59	50	109	0,8
Manta	56	31	87	0,7
Riobamba	49	44	93	0,7
Extranjero	47	0	47	0,4
Ibarra	44	58	102	0,8
Otros	317	556	873	6,6
Total	6.491	6.780	13.271	100,0

Fuente: CompuTrabajo, elaboración de los autores.

El salario asociado a la demanda laboral de una empresa es una variable sensible de presentar en las páginas web. Incluso, en algunos anuncios laborales es la empresa quién solicita al postulante que detalle el monto o rango de interés. La información del salario, es uno de los factores determinantes tanto para la oferta como para la demanda laboral de continuar en un proceso de selección.

Respecto a la información de salarios en este ejercicio, estos son presentados en rangos para facilitar su análisis. El primer intervalo se construyó en función del Salario Básico Unificado (SBU) vigente para 2016 que corresponde a US\$ 366 mensuales, en tanto que el último intervalo fue acotado en US\$2.000 y más, dado el reducido número de observaciones para este rango. De la Tabla 5 se desprende que, un 42% de plazas de trabajo no mencionan el salario y prefieren acordarlo con los candidatos al momento del reclutamiento, un 52% de plazas de trabajo ofrecen un salario menor a US\$900 y, llama la atención que cerca de un 10% de vacantes propongan salarios menores al SBU.

En octubre se aprecia una mayor oferta de vacantes con salarios menores a los US\$400. Un análisis con una mayor temporalidad permitirá conocer si esto es una coincidencia u obedece a algún fenómeno particular.

Tabla 5. Distribución de salarios

Rangos salariales	Sep	Oct	Total	%
0-366	554	705	1.259	9,5
366-400	920	1.222	2.142	16,1
400-500	654	590	1.244	9,4
500-600	493	530	1.023	7,7
600-700	321	377	698	5,3
700-800	124	95	219	1,7
800-900	171	79	250	1,9
900-1000	50	27	77	0,6
1000-1250	119	109	228	1,7
1250-1500	22	8	30	0,2
1500-1750	44	30	74	0,6
1750-2000	9	4	13	0,1
2000+	14	7	21	0,2
A convenir	2.855	2.767	5.622	42,4
No especifica	141	230	371	2,8
Total	6.491	6.780	13.271	100,0

Fuente: CompuTrabajo, elaboración de los autores.

Como se puede observar en la Tabla 6, la experiencia laboral es un elemento determinante para acceder a una plaza de trabajo. Apenas el 3% de vacantes no requieren ningún tipo de experiencia para ser satisfechas mientras que para el 85% de los puestos de trabajo se condiciona el tener tiempo de experiencia; el 12,7% de plazas no especifican la necesidad de contar o no con experiencia. De las vacantes que requieren experiencia, la mayoría solicitan tener al menos un año de experiencia (48%), seguido de las que necesitan al menos dos (21%) y tres años de experiencia (10%). Es fácil notar que existe una relación inversa entre años de experiencia y plazas de trabajo en los anuncios.

Tabla 6. Distribución de los años de experiencia

Experiencia	Sep	Oct	Total	%
0	192	164	356	2,7
1	3.063	3.306	6.369	48,0
2	1.351	1.423	2.774	20,9
3	635	734	1.369	10,3
4	147	93	240	1,8
5	209	174	383	2,9
6	12	12	24	0,2
7	4	5	9	0,1

8	27	3	30	0,2
10	10	24	34	0,3
No especifica	841	842	1.683	12,7
Total	6.491	6.780	13.271	100,0

Fuente: CompuTrabajo, elaboración de los autores.

Al analizar en conjunto la relación entre años de experiencia y salario propuesto, se puede observar que existe una correlación positiva. El 59% de las plazas de trabajo ofrecen menos de US\$ 700 y exigen un año de experiencia o menos. Las vacantes que ofertan US\$2.000 o más exigen tener al menos 3 años de experiencia (Tabla 7).

Tabla 7. Rangos salariales y años de experiencia.

Salario vs Exp	0	1	2	3	4	5+	Total %
0-366	0,20%	12,50%	1,10%	0,20%	0,30%	1,00%	15,30%
366-400	1,40%	20,40%	5,00%	0,70%	0,20%	0,20%	27,90%
400-500	1,80%	6,80%	5,50%	2,10%	0,20%	0,20%	16,60%
500-600	0,00%	8,40%	3,10%	1,50%	0,10%	0,10%	13,20%
600-700	0,70%	6,70%	1,00%	1,20%	0,20%	0,00%	9,80%
700-800	0,00%	1,10%	1,20%	0,70%	0,40%	0,40%	3,80%
800-900	0,00%	3,80%	0,90%	0,60%	0,10%	0,40%	5,80%
900-1000	0,40%	0,10%	0,40%	0,10%	0,00%	0,00%	1,00%
1000-1250	0,00%	0,80%	1,30%	0,80%	0,80%	0,20%	3,90%
1250-1500	0,00%	0,10%	0,40%	0,20%	0,00%	0,00%	0,70%
1500-1750	0,00%	0,10%	0,30%	0,50%	0,50%	0,20%	1,60%
1750-2000	0,00%	0,00%	0,00%	0,00%	0,00%	0,30%	0,30%
2000+	0,00%	0,00%	0,00%	0,10%	0,10%	0,20%	0,40%
Total %	4,50%	60,80%	20,20%	8,70%	2,90%	3,20%	100%

Fuente: CompuTrabajo, elaboración de los autores.

En lo que respecta a las características educativas de las vacantes, en la Tabla 8 se observa que en su mayoría (43%) requieren aspirantes que al menos hayan concluido el bachillerato, un 22% exigen título universitario y apenas un 1% demanda profesionales de cuarto nivel. Entre septiembre y octubre la demanda de bachilleres aumentó mientras que la de universitarios bajó en una proporción similar.

Esta mayor demanda de personal menos calificado puede responder a un fenómeno estacional, sin embargo, es todavía una hipótesis a ser comprobada⁹.

⁹ Se debe considerar que en el presente ejercicio el periodo de análisis es muy corto.(2 meses)

Para enfrentar algunas de estas debilidades se ha mencionado como alternativa el incluir en alguna de las encuestas dirigidas a empresas (o quizás diseñar una encuesta para el efecto) un conjunto de preguntas que permitan poner en contexto la información de las vacantes publicadas en internet. Por ejemplo, sería útil conocer qué porcentaje de las empresas usan este tipo de bolsas de empleo para promocionar sus necesidades de mano de obra, y algunos datos adicionales para caracterizar estas prácticas y sus tendencias.

Aún con las debilidades conocidas, es probable que las estadísticas de las vacantes publicadas en internet sean una forma útil y práctica de medir el pulso de la economía. Sin embargo, la evaluación de la relación que puede existir entre el número de vacantes publicadas y la actividad económica es algo que sólo puede hacerse una vez que se cuente con un rango temporal de la información lo suficientemente extenso. De cualquier forma, conocer en tiempo real las demandas de las empresas, respecto del personal a contratar, puede proveer de información útil sobre el tipo de perfiles de trabajadores que se requieren.

Dado que este artículo se basó en el análisis de dos meses de referencia, se recomienda realizar la investigación sobre una temporalidad mayor con la finalidad de poder identificar patrones y fluctuaciones en las variables consideradas.

7. Referencias

Antweiler, W. & Murray, F., 2004. Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards.. *Journal of Finance*, 1259-94(59), p. 1.

Askatas, N. & Zimmermann, K. F., 2015. The Internet as a data source for advancement in social sciences. *Institute for the Study of Labor*, Issue 8899, p. 3.

Azar, O. H., 2007. The Slowdown in First-Response Times of Economics Journals: Can it Be Beneficial?. *Economic Inquiry*, 179-87(45), p. 1.

Bajari, P. & Hortacsu, A., 2003. The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions. *RAND Journal of Economics*, 329-55(34), p. 1.

Benítez, D.; Espinoza, B.; Grijalva, A.; Rivadeneira, Ana.; Oviedo, Ana María., 2016. Using administrative

records to study employer-employee dynamics: The case of Ecuador's Laboratory of Labor and Business Dynamics. *Working Paper*, Issue 1, pp. 2-4.

Cárdenas, J., Guataquí, J. C. & Montaña, J., 2015. Metodología para el análisis de demanda laboral mediante datos de Internet: el caso colombiano.

Carnevale, A. P., Jayasundera, T. & Repnikov, D., 2014. Understanding Online Job Ads Data. A Technical Report. *Georgetown University*, p. 17.

Castrillo-Fernández, O., 2015. Web Scraping: Applications and Tools. *European Public Sector Information Platform*, Issue 2015/10, pp. 1-15.

Cavallo, A., 2009. Scraped Data and Sticky Prices: Frequency, Hazards, and Synchronisation. *Harvard University Job Market Paper*, p. 5.

Cavallo, A., 2012. Online and official price indexes: measuring Argentina's inflation. *Massachusetts Institute of Technology*, p. 9.

Edelman, B., 2012. Using Internet Data for Economic Research. *American Economic Association*, 26(2), p. 192.

INEC, Instituto Nacional de Estadísticas y Censos., 2016. Metodología para la medición del empleo en el Ecuador. pp. 8-10.

Kuhn, P. & Mansour, H., 2011. Is Internet Job Search Still Ineffective?. *Institute for the Study of Labor (IZA) DP*, Issue 5955, p. 1.

Kurekova, L. M., Miroslav, B. & Anna, T.-T., 2015. Using online vacancies and web surveys to analyse the labour market: a methodological inquiry. *IZA Journal of Labor Economics*, Volumen 4:18, pp. 2-3.

Reimsbasch-Kounatze, C., 2015. The Proliferation of "Big Data" and Implications for Official Statistics and Statistical Agencies. A preliminary analysis.. *OECD Digital Economy Papers*, Issue 245, p. 15.

Seamans, R. & Zhu, F., 2010. Technology Shocks in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers. *NET Institute Working Paper*, 10(11), p. 1.

Stefanik, M., 2012. Internet job search data as a possible source of information on skills demand (with results for Slovak university graduates). *Building on skills Forecasts-Comparing Methods and Applications*, p. 1.