

Big Data para la medición de la economía de Internet

Aline Visconti Rodrigues

10/12/2020

Objetivo del proyecto



Medir la economía digital utilizando herramientas y técnicas de Big Data y su combinación con estadísticas oficiales.

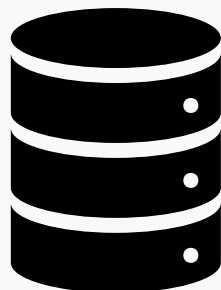


NACIONES UNIDAS

CEPAL

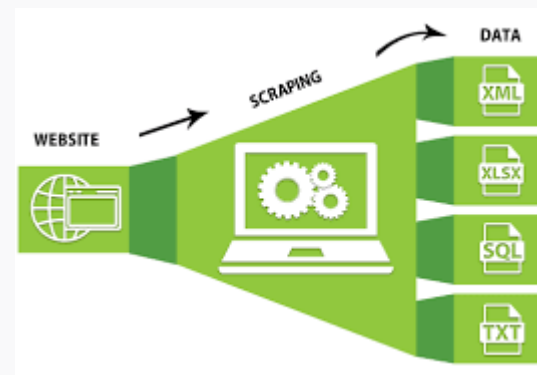
¿De qué manera?

Registros administrativos
y bases de datos de estadísticas
oficiales

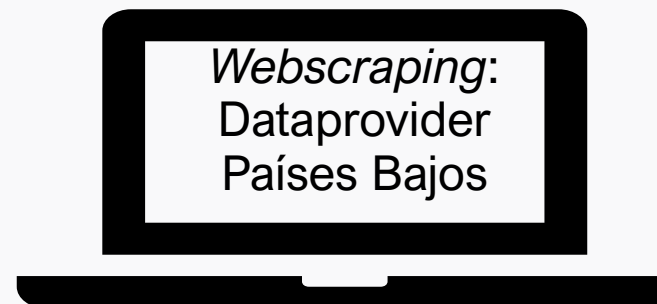


X

Datos recolectados
mediante técnicas de *web
scraping*



¿Cuáles fueron las fuentes en caso de Brasil?



¿Cómo fue el ejercicio? Basado en la experiencia de los Países Bajos

Measuring the internet economy in The Netherlands: a big data analysis, (Oostrom L., y otros, 2016)

Primera etapa



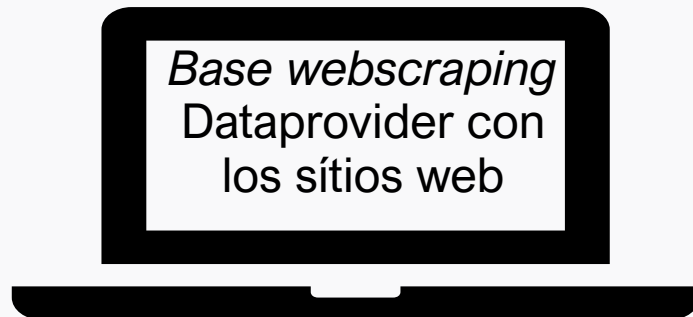
Conceptualizar la economía de internet mediante el uso de los sitios web por empresa. Para ello, fueron creadas categorías para cada nivel de presencia en la web.

Segunda etapa



Recopilación y procesamiento de datos: base de datos con datos de sitios web corporativos obtenidos por Dataprovider y registros oficiales de IBGE y de Nic.br.

Primer paso



+

Data mining



Palabras clave por categoría

Continuación del primer paso

Seis categorías

A - Sin presencia

no webs

Sin sitios web vinculados

B1 - Presencia Passive



Peluquería

Sitios que no fueron categorizados en las categorías anteriores sin alguna interacción

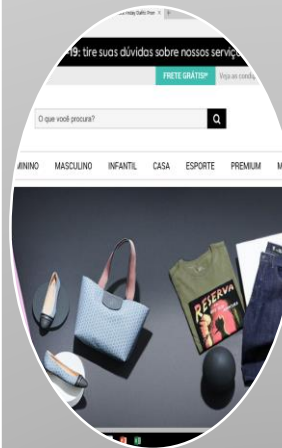
B2 - Presencia Active



Hoteles

Sitios que no fueron categorizados en las categorías anteriores con alguna interacción

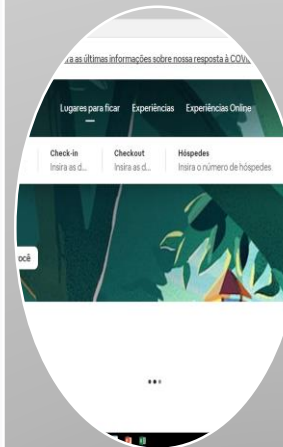
C-Ecommerce



Dafiti.com.br

Ecommerce

D- Online Services



Booking, Airbnb

Sitios de empresas que su renta es a través de la internet

E - Internet related ICT

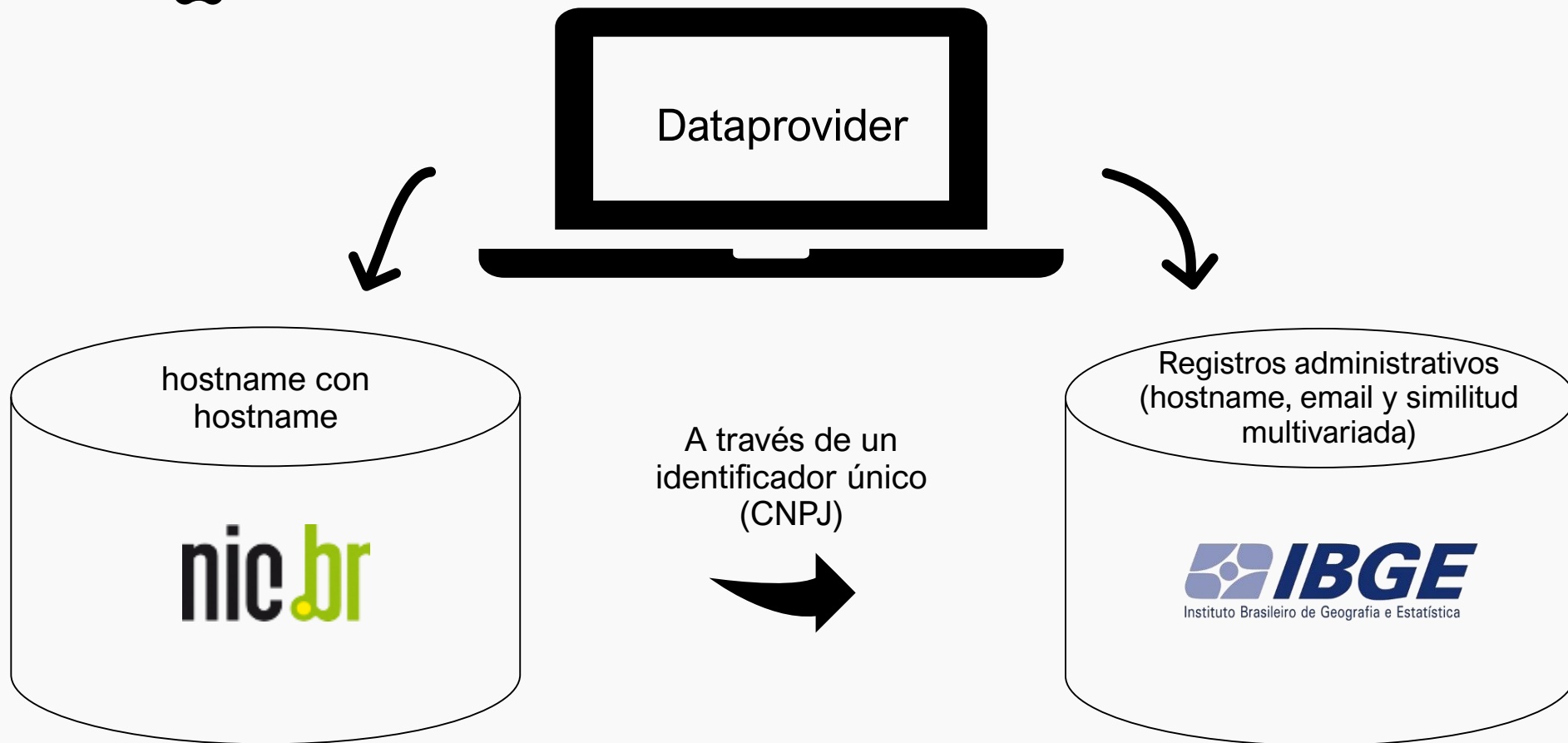


Webdesign, Hosting

Sitios de empresas que su renta es generada por la internet

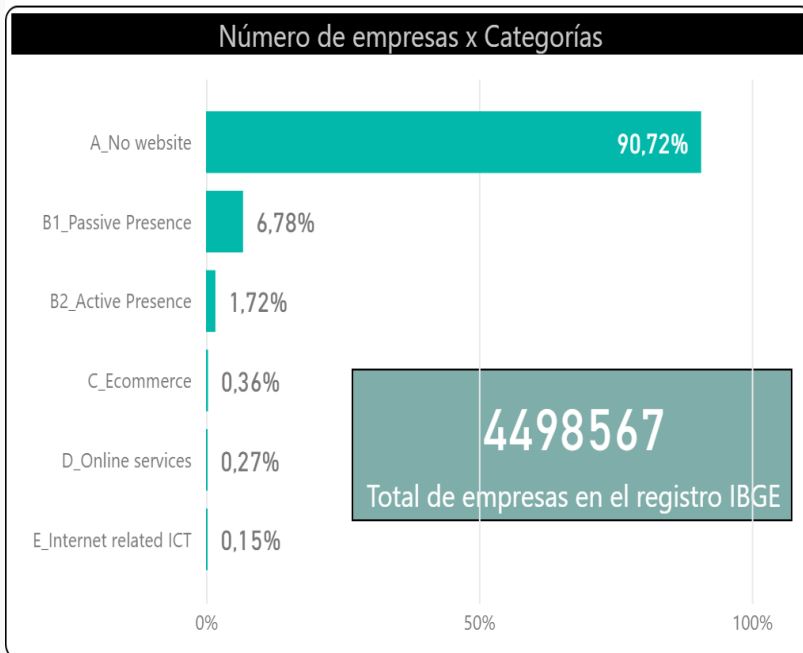
Segundo passo

Vinculación de las bases



Ultima etapa

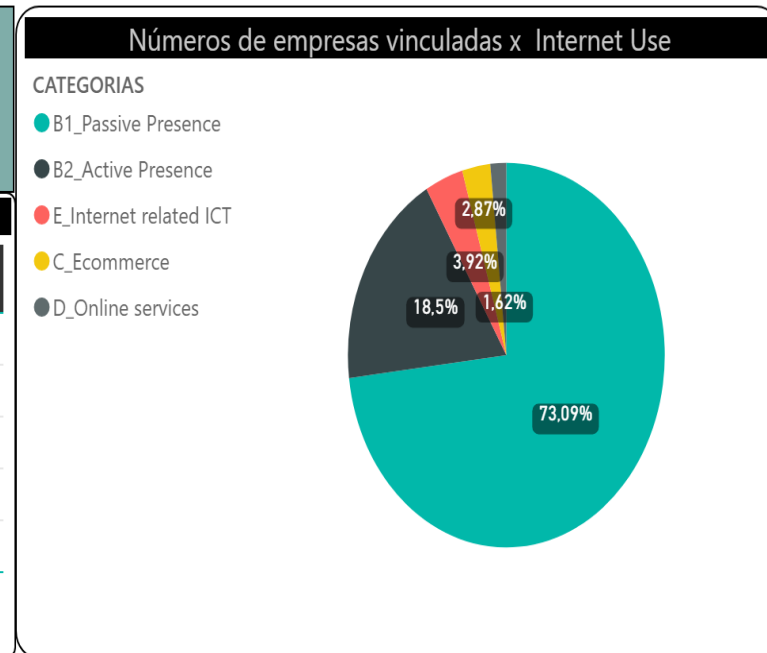
Análisis demográfico de la economía de internet - cálculo de indicadores



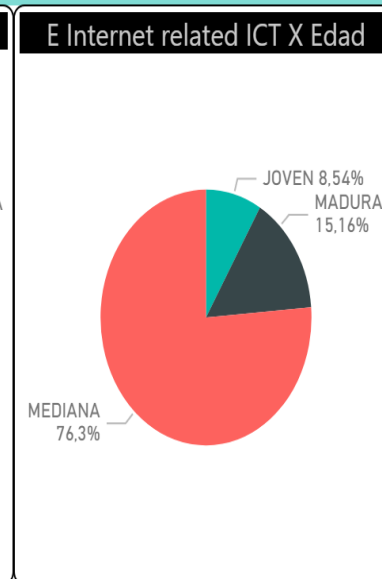
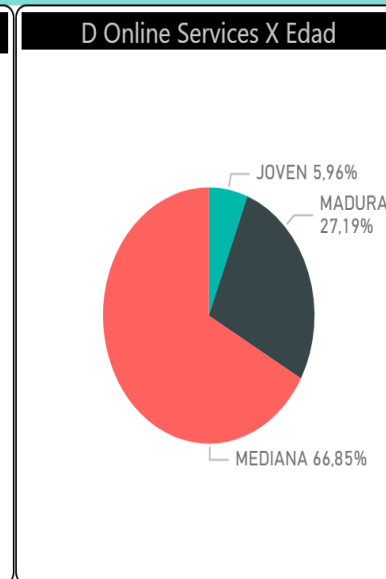
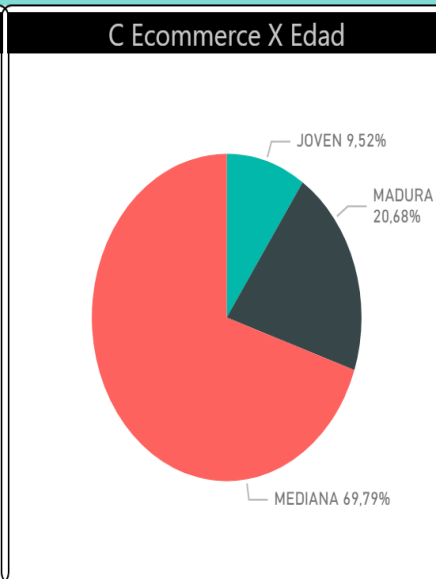
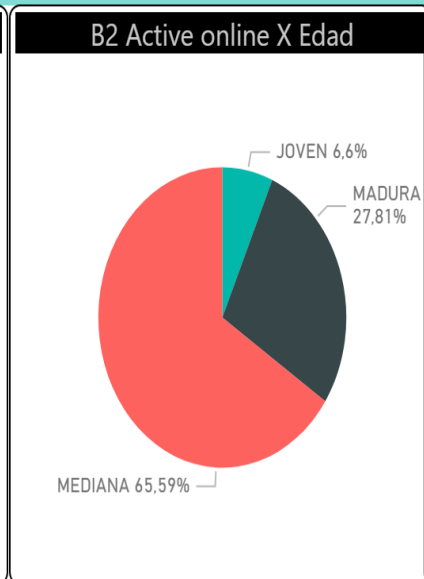
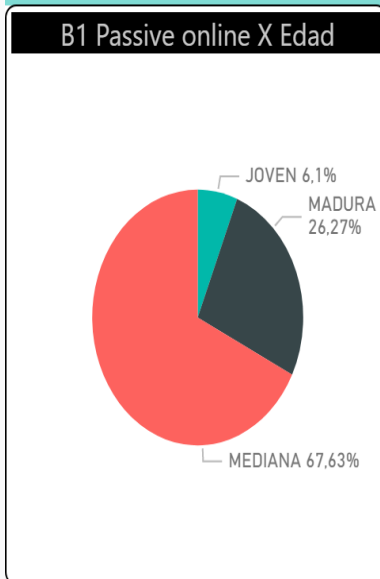
417360
Total de empresas vinculadas

Empresas vinculadas por categorías

Categorías	Empresas
B1_Passive Presence	305053
B2_Active Presence	77206
E_Internet related ICT	16354
C_Ecommerce	11990
D_Online services	6757
Total	417360

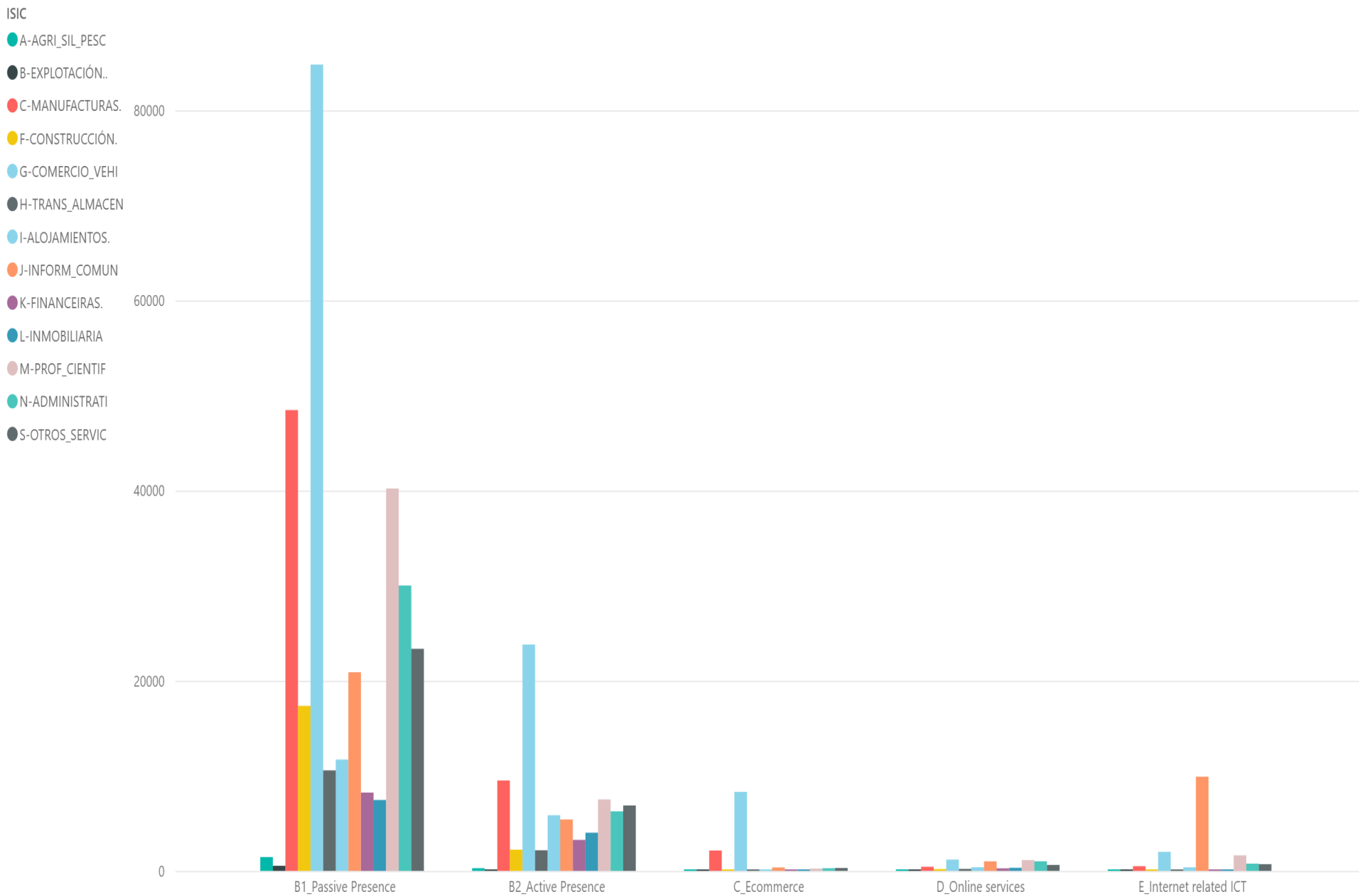


Joven - Empresa hasta 4 años de edad. /Mediana - 5 a 24 años de edad. / Madura - más o igual a 25 años de edad
Las categorías Ecommerce e Internet related ICT presentan las mayores tasas de empresas jóvenes (9,5% y 8,5%).



Las empresas del sector de Comercio son la mayoría en todas las categorías, excepto para la categoría Internet related ICT, en la cual las empresas del sector de Información y Comunicación tiene predominio.

Número de empresas vinculadas por Categorías X ISIC



Aprendizajes



Innovación – mezcla de estadísticas oficiales y datos no estructurados



El proceso de limpieza de datos es esencial. Big data no sustituye las estadísticas oficiales.



Cooperación entre los actores involucrados – IBGE, CEPAL, Nic.br.



La tardanza por parte de la administración del proyecto en IBGE.



Interacción dentro de IBGE



Coordinación de Encuestas Estructurales y Especiales en Empresas
Coordinación de Registros y Clasificaciones
Coordinación de Metodología
Coordinación de Índice de Precios.



Capacitación en técnicas de Big Data – apoyo de CEPAL.

¡¡Muchas gracias!!

Aline Visconti Rodrigues

**Coordinación de Encuestas Estructurales y Especiales en
Empresas - COSEC**

aline.rodrigues@ibge.gov.br