

Análisis Encuesta Origen Destino

Hipótesis 1: *"La mujer realiza más viajes a pie que el hombre. La diferencia es inversamente proporcional al nivel de ingreso de las comunas."*

Hipótesis 2: *"La mujer realiza más viajes en bicicleta que el hombre. La diferencia es inversamente proporcional al nivel de ingreso de las comunas."*

Para refutar o apoyar estas hipótesis, el objetivo será buscar una tabla que tenga, por comuna, la cantidad de viajes en el modo buscado (caminata o bicicleta), tanto para hombres como para mujeres, el total de viajes en dicho modo, y el ingreso de dicha comuna.

Cuando se intenta buscar el ingreso por comuna, los supuestos a realizar en base a los datos de la Encuesta Origen Destino (en adelante EOD), dejan mucho lugar para conclusiones equívocas ya que todo tipo de supuesto posible con los datos presentes tendrá un sesgo muy grande.

Dentro de las posibilidades estudiadas estuvieron: 1) considerar el ingreso de todas las personas que realizan viajes que se originan en la comuna; 2) considerar la mitad del ingreso de todas las personas que entran y salen de la comuna. Sin embargo, estas posibilidades no eran buenas opciones ya que es posible verificar en la EOD que muchos viajes son de ida y vuelta, por lo que reportarían ingresos de una persona a ambas comunas, y nada asegura que realizar viajes desde una comuna significa que la persona vive en ella, etc.

En base a eso, para poder llevar a cabo el análisis se obtuvo una base de datos en formato *csv* (que fue formateada para luego poder ser importada) desde el sitio web del Observatorio Urbano del MINVU (<http://www.observatoriourbano.cl/indurb/indicadores.asp>), que señala los ingresos medios de los hogares por comuna para el año 2009.

En base a eso, para fabricar la tabla objetivo antes mencionada, después de importar debidamente los archivos *csv*, la consulta realizada fue:

```
CREATE VIEW totalcaminata AS
SELECT
    c.comuna,
    COUNT(case when p.sexo="1" then m.id end) as hombre,
    COUNT(case when p.sexo="2" then m.id end) as mujer,
    COUNT(m.id) as total,
    CAST(i.ingreso as integer) as ingreso
FROM
    viajes AS v, personas AS p, comuna AS c, modo AS m, ingreso AS i
WHERE
    v.persona = p.persona AND v.comunaorigen = c.id AND
    m.id = v.modoadgregado AND i.comuna = c.comuna AND m.id = '17'
GROUP BY
    c.id
```

De la cual se fabrica otra consulta (que podría haber sido realizada en la misma consulta, pero por un tema de orden y entendimiento se realiza separadamente), la cual tiene por columnas a la comuna, la diferencia porcentual entre hombres y mujeres y el ingreso medio de los hogares de la comuna:

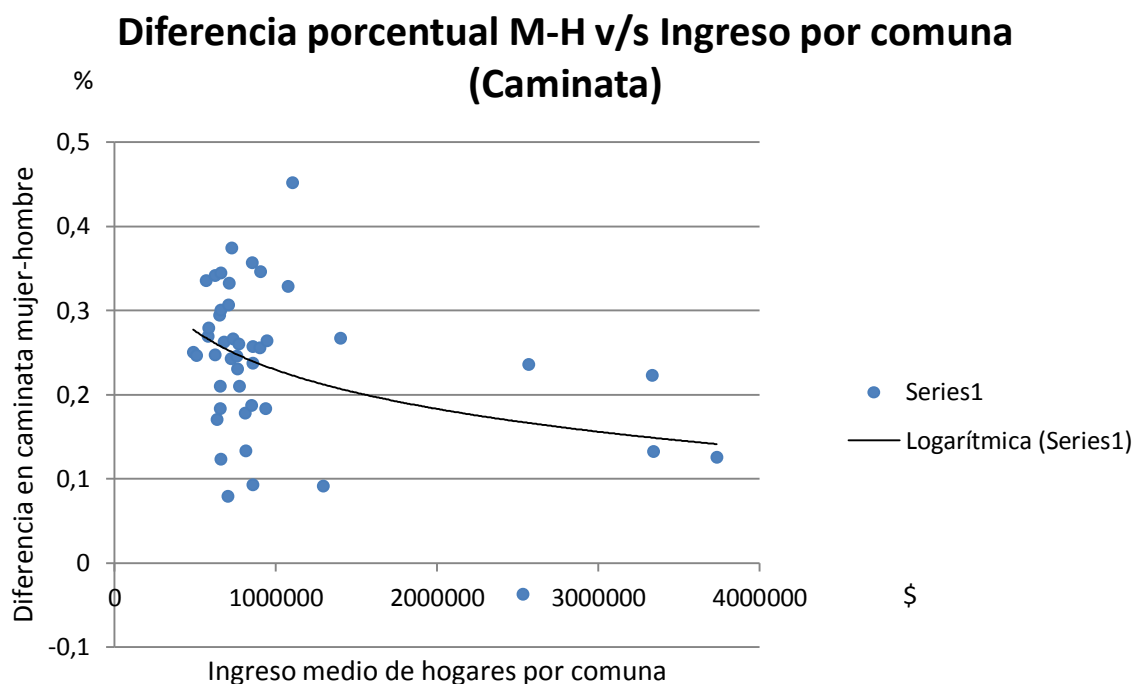
```
SELECT
    t.comuna,
    (mujer*1.0)/total - (hombre*1.0)/total AS diferencia,
    ingreso
FROM
    totalcaminata AS t
ORDER BY
    ingreso
```

Se procede análogamente para los viajes en bicicleta, con la diferencia de que se restaron las mujeres a los hombres para tener una diferencia positiva, y el id para ese modo fue 18 en vez de 17.

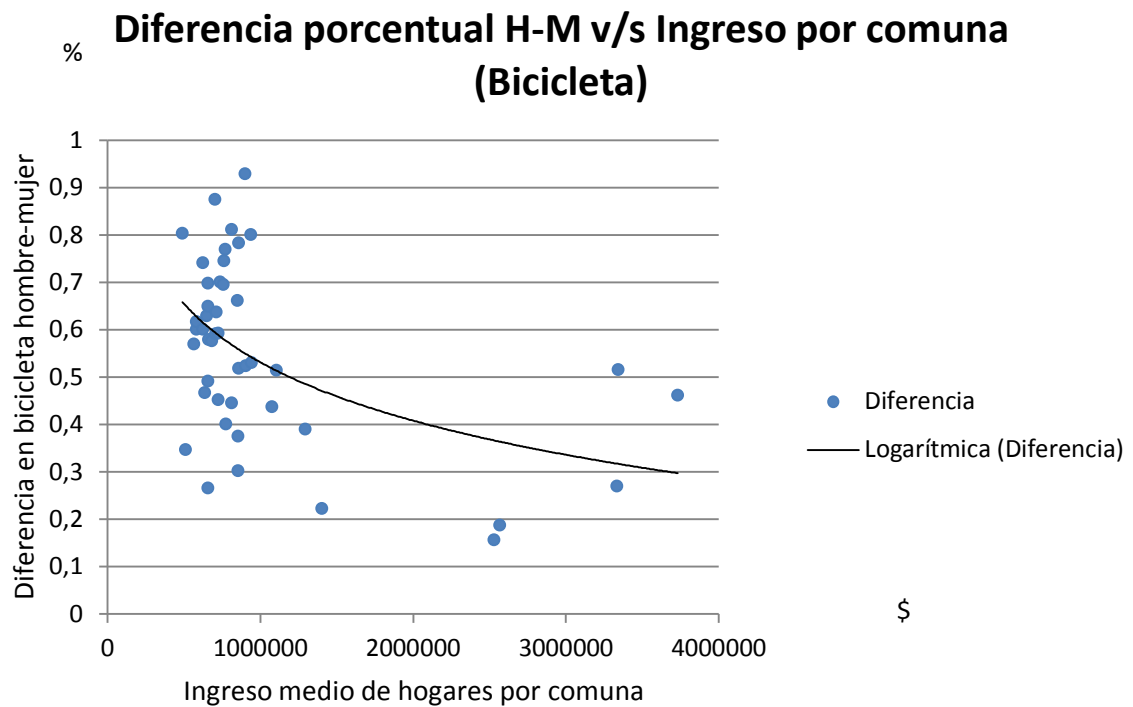
Parte del resultado final para la última consulta en el caso de la caminata es:

Comuna	Ingreso	Diferencia
La Pintana	489010	0,25
Isla de Maipo	508522	0,2462
La Granja	565953	0,3349
...
La Reina	2566862	0,2356
Vitacura	3332463	0,2222
Lo Barnechea	3341069	0,1317
Las Condes	3733792	0,1254

Y el gráfico corresponde a:



Para el caso de las bicicletas el gráfico es este:



En ambos casos se observa una relación inversa entre la diferencia porcentual y el ingreso medio de hogares, pero esta relación es muy débil y difícilmente constituye argumento suficiente para afirmar la hipótesis. En el segundo caso quien anda más en bicicleta es el hombre pero igualmente existe una disminución en la brecha en las comunas de mayores ingresos.

Para que sea considerada como una proporción inversa, la curva debiese ser similar a la curva de la función $f(x) = \frac{1}{x}$, sin embargo, puede observar cierta similitud en el ajuste logarítmico presentado en el gráfico, pero claramente no sigue a los datos de manera precisa.

Se concluye que no es posible confirmar la hipótesis con seguridad para dicha encuesta pero que si existen indicios de la relación inversa propuesta.