



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE  
INSTITUTO DE INGENIERÍA MATEMÁTICA Y COMPUTACIONAL  
LIC. EN ING. EN CIENCIA DE DATOS  
IMT-2200

# **La relación entre el transporte público en la región metropolitana y los diferentes índices de su población**

---

*Propuesta*

**Grupo 8**

**Integrantes:**

Jeferson Alzate, Joaquín Bächler, Matias Morales, Franco Moya, Gabriel Vergel

**Fecha de entrega:** 8 de octubre de 2024



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE  
INSTITUTO DE INGENIERÍA MATEMÁTICA Y COMPUTACIONAL  
LIC. EN ING. EN CIENCIA DE DATOS  
IMT-2200

## Índice de contenido

<b>1. Contexto y Motivación</b>	<b>2</b>
<b>2. Objetivos</b>	<b>2</b>
<b>3. Datos</b>	<b>4</b>
<b>4. Preguntas de Investigación</b>	<b>6</b>
<b>5. Diseño tentativo</b>	<b>7</b>
5.1. Procesos Estadísticos . . . . .	7



## 1. Contexto y Motivación

Dado que somos usuarios cotidianos del servicio de transporte público, es natural que tengamos un interés en él. A diario nos encontramos con vagones (o buses) llenos, vacíos o normales, y surgen las dudas; ¿Qué hay de los usuarios de esos vagones?, ¿Hay algún patrón?, ¿Alguna tendencia?, ¿Siempre ha sido así?

Son preguntas que uno podría hacerse, ya sea por curiosidad, por planear mejor un viaje, por entender mejor, o por lo que sea. La cuestión es que hay algo interesante que puede ser investigado.

Pensamos que a través de este proceso podremos responder alguna de estas preguntas, y si no es así, quizá alguna nueva intriga que surja durante el proceso de recopilación o trabajo con los datos.

Lo bueno de esta iniciativa es la utilidad que tiene para gran parte de la población, pues una significativa parte de ella hace uso del sistema de transporte público. Por otro lado, entendiendo que el órgano que regula este sistema debe estar informado acerca de su propio funcionamiento, podríamos decir que la información (más que la información, las conclusiones) podrían de servirle en algún grado.

## 2. Objetivos

Para marcar un enfoque más claro y acotado, decidimos concentrarnos en hallar qué relaciones existen entre los factores socioeconómicos de la población y el transporte público, todo esto solo dentro de la región metropolitana. El objetivo principal es entender como se comporta la riqueza de las personas con el transporte público en la región metropolitana.

En base a lo anterior, la pregunta principal de este proyecto será ¿Cómo podemos cuantificar la influencia del estado del transporte público sobre el estatus económico de las personas? Creemos que hallar formas de cuantificar la relación *transporte-riqueza*, puede aportar a un profundo entendimiento de la dinámica entre estos dos ámbitos.

A partir del proceso que nos proponemos llevar acabo podremos obtener información acerca de :

- Disponibilidad de transporte
- Cantidad de uso del transporte
- Cambios en la estructura socioeconómica en base al transporte público
- Preferencias de la población en el transporte



- Factores demográficos y el uso del transporte
- Influencia de fenómenos inesperados dentro del sistema de transporte (como el COVID-19)
- Proyecciones sobre el comportamiento del transporte

Mediante este trabajo, pretendemos entregar una mayor claridad a la hora de tomar decisiones que involucren cambios en el estado del transporte, que a través de un panorama más transparente acerca del cómo estas decisiones pueden afectar o beneficiar a las personas, puedan efectuarse resoluciones más positivas. También pretendemos que el esclarecimiento de esta relación *transporte-riqueza* sirva para mejorar la comprensión general de cómo los diversos factores de la sociedad pueden influir en nuestra vida y nuestro futuro.

Creemos que este proyecto puede ser de utilidad principalmente para los usuarios comunes del metro y los buses, porque son quienes más provecho podrían obtener del conocimiento estadístico que recopilamos, aplicándolo en su toma de decisiones diarias. Sin embargo también puede llegar a ser útil para otras personas que busquen desarrollar un trabajo investigativo acerca de temas relacionados al nuestro, o incluso podría llegar a ser de alguna utilidad para aquellos que proponen y efectúan nuevos proyectos sobre el estado de la red de transporte.

Recordando que el problema fundamental de esta investigación es entender *la relación entre el transporte público y la población de Santiago*, y a partir de los puntos expuestos más arriba, podríamos decir que una pregunta amplia que abarca de manera general todos estos aspectos sería : **¿Qué historia podemos contar?**. En otras palabras, y respondiendo lo anterior, ¿Nos encontramos con patrones, con fallas, con implementaciones satisfactorias, con posibles mejoras, proyecciones, etc.?



### 3. Datos

Para llevar a cabo este proyecto, dispondremos, principalmente de 9 (o más) fuentes de datos.

1. [Informes del Directorio de Transporte Público](#)
2. [Estadísticas del registro social de hogares](#)
3. [Información recopilada hasta el último censo \(disponible\) 2017](#)
4. [API openstreetmap \(mapa santiago y recorridos de transporte\)](#)
5. [Datos Abiertos BIDAT \(RSH\)](#)
6. [ADIS \(Datos poblacionales\)](#)
7. [Permisos de circulación \(INE\)](#)
8. [Reportes Comunales \(BCN\)](#)
9. [Reporte de Movilidad Ingeniería UC](#)

*El desarrollo y manejo de estos datos se puede ver en nuestro [repositorio de Github](#).*

En el primero podremos encontrar diferentes índices, estadísticas y registros que pertenecen al **DTP**. Lo bueno de esta base, es que los datos son estructurados. Podemos encontrar las siguientes tablas de información relacionadas con los índices de evasión a lo largo del tiempo, cantidad de usuarios por servicio, precios de los pasajes a lo largo del tiempo, horarios con mayor demanda del servicio, sesgo etario (a partir de la implementación de los pases de tercera edad) , etc.

Por otro los datos del **INE** cuentan con un abanico más amplio de opciones, pero por lo mismo es una oportunidad para poder hacer y/o filtrar los datos de manera que nos sea útiles. En este caso haremos consultas relativas a diferentes índices considerados en los *censos* para luego poder relacionarlas con la data obtenida de la fuente anterior.

Por lo mismo el **registro social de hogares** nos brinda información necesaria para cruzarla con la anterior, pudiendo hacer visualizaciones sobre la relación uso del servicio - situación socioeconomica. Estos cruces pueden ser observados particularmente (líneas y/o recorridos específicos), como globalmente (mapa de la región).

Por esto último podemos integrar la información de la API (openstreetmap) para visualizar geográficamente (espacialmente) distintas variables y cómo estas se relacionan con el servicio de transporte.



También estaremos ocupando el servicio de **RedAtam Webserver**, que provee una solución para las consultas a datos de esta índole. Funciona como una especie de **query** que devuelve un archivo excel con la información solicitada.

Además encontramos una serie de otras fuentes que podemos interconectar con las anteriores, de manera que podamos tener información más robusta o bien rectificar ciertos datos (almacenados en diferentes lugares).



## 4. Preguntas de Investigación

Hay muchas cosas que podríamos preguntarnos a partir de los datos que utilizaremos.

- ¿Qué relación existe entre el volumen de los sectores etarios y el consumo de transporte público?
- ¿Cómo afectó la pandemia (COVID-19) el uso de este servicio?
- ¿Cómo se relaciona el mercado automotriz con el uso del transporte?
- ¿Qué relación existe entre la disponibilidad del transporte público y la riqueza de las personas en cada comuna de la región?
- ¿Cómo se relaciona la densidad demográfica con el uso del transporte?
- ¿Cómo se relaciona el cambio en el valor del servicio con el uso de este ?
- ¿Cómo afectó al metro la creación de las dos nuevas líneas? ¿Se logró contener de mejor manera la afluencia?
- ¿Qué recorridos presentan más uso, qué podemos concluir de esto ?
- ¿Cómo vemos la proyección del uso del usuario respecto a este servicio? (aumenta, se mantiene, baja)

Trataremos de dar respuesta a cada una de estas preguntas en base a mediciones y resultados estadísticos cuantificables.



## 5. Diseño tentativo

En primer lugar pensamos que lo mejor es poder ordenar (pasar a csv, json y luego pandas) toda la información que podamos recolectar, de manera que a la hora de responder preguntas podamos trabajar más fácilmente con los datos que tenemos. De esta forma la información se mantendrá intacta y podremos hacer los procesos necesarios en archivos diferentes según la pregunta a responder.

Ej :

```
1  from pasajeros import cantidad_usuarios_metro
2
3  cantidad_usuarios_metro.head()
4
5..
6
7  c_2010 = cantidad_usuarios_metro[(cantidad_usuarios_metro['ano'] >= 2000) & (cantidad_usuarios_metro
   [ 'ano' ] <= 2010)]
8
9  ...
10
11 print(f'Promedio de pasajeros entre 2000 y 2010 : {prom20_10}')
```

### 5.1. Procesos Estadísticos

Dado que tenemos una gran cantidad de datos, de diferente índole, nos encontramos con un abanico de posibilidades, en cuanto a procedimientos. En primer lugar pensamos que una regresión lineal puede ser una buena opción para entender (y proyectar) el comportamiento del servicio entero, como de sus diferentes ramas (en cantidad de pasajeros, líneas, precios, ingresos, etc) .

Esta herramienta puede de ser de gran utilidad a la hora de observar diferentes correlaciones entre los índices utilizados, permitiendonos ver si es que realmente existe una tendencia para luego proyectarla a lo largo del tiempo.

También nos parece importante implementar el uso de la correlación (en forma matricial) de diferentes índices. Lo cual pensamos también nos aportara mayor claridad sobre las relaciones socio-economicas y el uso del transporte, así como las demograficas y etarias.

Pensamos que visualizaciones de mapa (colormap) serán útiles para este caso, donde podemos indicar la variabilidad de un índice (eg: demografico) a través del degradez, superponiendo el esqueleto del





PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE  
INSTITUTO DE INGENIERÍA MATEMÁTICA Y COMPUTACIONAL  
LIC. EN ING. EN CIENCIA DE DATOS  
IMT-2200

metro/red bus , y observar que fenomenos ocurren, y entender el problema de desde otra perspectiva (espacial).

De esta manera pretendemos poder abordar, entonces, el proyecto desde diferentes aristas, pero todas apuntando a lo mismo : **entender cómo se relacionan los índices de la población de la R.M con el uso del transporte público.**