

Jonathan Déniz Orosa

*Modelos de demanda directa y análisis  
de datos de la movilidad de pasajeros*

Direct demand models and passenger mobility  
data analysis

Trabajo Fin de Grado  
Grado en Matemáticas  
La Laguna, Junio de 2022

DIRIGIDO POR  
*Carlos Pérez González*

*Carlos Pérez González*  
*Matemáticas, Estadística e*  
*Investigación Operativa*  
*Universidad de La Laguna*  
*38200 La Laguna, Tenerife*

---

## Agradecimientos

Después de un largo camino, he de agradecer a muchas personas que durante el trayecto me han ayuda a conseguir mi gran objetivo: ser Matemático, y poder así entender una ciencia tan pura, compleja, que abarca tantos campos, y que incluso en sus elementos mas intuitivos te ayudan a comprender el por qué de las cosas.

Agradecer a cada uno de los profesores que he tenido, he aprendido de todos, adquiriendo madurez, autonomía, pensamiento crítico, entre otras muchas cosas. Sobretudo, agradecer a Juan Cristo Luis Jorge, el cuál un día nos dijo una frase surgida de su experiencia y madurez: “El trabajo duro siempre supera al talento”. Esta frase es la que he llevado grabada a fuego en mí desde el principio, dedicando todo mi esfuerzo y tiempo en mejorar y obtener los mejores resultados.

Agradecer a mis padres, los cuales me han inculcado todo y cada uno de los valores humanos que poseo, al igual que me han prestado todo su apoyo para que yo cumpliera mi sueño de ser Matemático. Aunque todo el mundo deba decirlo, yo lo pienso de verdad, y es que tengo la mejor familia que una persona podría tener. Gracias a ellos he vivido momentos tan únicos e irrepetibles, dandome todo lo que podían y tenían para que yo fuera feliz. Mamá, papá, os quiero en el alma.

A mi pareja Andrea, por ser mi sustento en los buenos y malos momentos, y no dejarme venir abajo nunca, valorandome como nadie. Siendo la única persona que me conoce a la perfección, y que solo ella sabía sacarme una sonrisa incluso cuando mas negro veía todo. Eres la persona más pura y talentosa que he conocido y conozco. No sabes la suerte que tengo de tenerte a mi lado.

A mis dos grandes amigos que me llevo de la carrera, Jorge y Fran, con los cuáles he formado una amistad pura y que perdurará muchos años, que han sufrido igual que yo, y con los que he reído, llorado, debatido sobre las matemáticas, y no tan matemáticas (dichosa Taylor Swift). Espero seguir teniendo con

ustedes nuestras famosas tarde de marujeo con café de por medio.

Agradecer a mi profesora Nana de mi etapa en el instituto, la cuál se convirtió en una madre para mí. A mi profesor Acosta, el artífice de que me enamorara de las matemáticas y decidiera dar el paso a estudiarla. También agradecerle a Antonio Lorenzo Bonilla, mi tutor de POAT, siempre interesado en mi progreso a lo largo de la carrera y el cuál siempre me ha dado consejos valiosos para lograr los mejores resultados.

Por último pero no menos importante, agradecerle a Carlos Pérez González por ser mi tutor de TFG, haberse interesado e inmerso en el trabajo con las mismas ganas y pasión que lo he hecho yo. Incontables horas de trabajo y tutorías, con problemas externos a nosotros que nos han ido retrasando, pero no han logrado quitarnos las ganas e ilusión de seguir adelante con este maravilloso trabajo.

Podría seguir agradeciendo a muchas personas más, familiares, amigos, y todos aquellos que me han ayudado a ser quien soy. Gracias, de corazón, a todos.

Jonathan Déniz Orosa  
La Laguna, 6 de julio de 2022

---

## Resumen • Abstract

### *Resumen*

---

*En el proceso tradicional de planificación del transporte y vialidad (Generación, Distribución, Repartición Modal y Asignación), se utiliza la fase de distribución para repartir los viajes que se producen entre las diferentes zonas que conforman una región, en este caso, la isla de Tenerife. En el presente trabajo se describen, tanto el marco teórico como la aplicación del Modelo Matemático de Gravedad utilizado para la distribución de viajes, así como sus alcances y limitaciones. Este modelo se calibrará a partir de los datos obtenidos por la entidad TITSA, con los orígenes y destinos de los viajes realizados por 717.944 pasajeros a lo largo de 9 días, previo análisis espacial y estadístico del conjunto de datos, utilizando el software QGIS y el entorno y lenguaje de programación estadístico R.*

**Palabras clave:** *Repartir – Modelo de Gravedad – TITSA – QGIS – Programación en R*

### *Abstract*

---

*In the traditional process of transport and road planning (Generation, Distribution, Modal Distribution and Allocation), the distribution phase is used to distribute the trips that are produced between the different areas that make up a region, in this case, the island of Tenerife. This paper describes both the theoretical framework and the application of the Mathematical Gravity Model used for trip distribution, as well as its scope and limitations. This model will be calibrated on the basis of the data obtained by TITSA, with the origins and destinations of the journeys made by 717,944 passengers over 9 days, after spatial and statistical analysis of the data set, using the software QGIS and the R statistical programming language and environment.*

**Keywords:** *Distribute – Gravity Model – TITSA – QGIS – R programming*



---

# Contenido

<b>Agradecimientos</b> .....	III
<b>Resumen/Abstract</b> .....	V
<b>Introducción</b> .....	IX
<b>1. Planificación y modelización del transporte</b> .....	1
1.1. Características de los problemas de transporte .....	2
1.1.1. Demanda de transporte .....	2
1.1.2. Oferta de transporte .....	4
1.1.3. Planificación del transporte .....	5
1.2. Elección de enfoques de modelización .....	5
1.3. Cuestiones de la modelización del transporte .....	6
1.3.1. Especificaciones del modelo .....	6
1.3.2. Calibración, validación y uso del modelo .....	6
<b>2. La estructura del modelo clásico de transporte</b> .....	9
2.1. Modelización de la generación de viajes .....	11
2.1.1. Definiciones básicas .....	11
2.1.2. Caracterización de los viajes .....	13
2.1.3. Factores que afectan a la generación de viajes .....	14
2.2. Modelización de la distribución de viajes: <i>Modelo de Gravedad</i> ...	14
2.2.1. Definiciones y notación .....	15
2.2.2. Generación del Modelo de Gravedad .....	18
<b>3. Análisis del flujo de pasajeros en Tenerife</b> .....	23
3.1. Fuente de datos .....	23
3.1.1. Uso de software QGIS .....	25
3.1.2. Archivos SHP .....	27
3.2. Exploración del conjunto de datos en RStudio .....	30
3.2.1. Introducción al análisis descriptivo .....	30

3.3. Primer caso de estudio: Santa Cruz de Tenerife .....	33
3.4. Segundo caso de estudio: San Cristobal de La Laguna .....	36
3.5. Estimación de pasajeros mediante modelo de gravedad .....	39
<b>A. Apéndice</b> .....	47
A.1. Código en lenguaje R para el análisis del conjunto de datos proporcionados por TITSA .....	47
<b>Bibliografía</b> .....	49
<b>Poster</b> .....	51

---

## Introducción

La modelización de transporte nos permite estimar los flujos de pasajeros o vehículos que habrán en una red de transporte en cada uno de los modos considerados para escenarios futuros.

Con esto, la *demanda de transporte de pasajeros* es un indicador de fuerza motriz, es decir, aquellos representativos de las actividades humanas, procesos de consumo o producción, que impactan en los objetivos de un desarrollo sostenible. Estos indicadores pueden ser de gran importancia para regular la demanda de transporte de pasajeros y fomentar modos de transporte específicos. El desglose de la demanda de transporte de pasajeros por modos además ayuda a evaluar la eficacia de las medidas de respuesta.

A lo largo de este trabajo se intentará entrever aquellos aspectos más importantes o a tener en cuenta en la modelización y la planificación del transporte, así como el nivel de detalle necesario de los mismos, intentando contribuir en este campo del conocimiento siguiendo la línea planteada en [1]. Se responderán a preguntas surgidas desde el comienzo del desarrollo de la demanda de transporte, como pueden ser:

- ¿Cuántos viajes realizará una persona, con un determinado objetivo, a lo largo de un tiempo estipulado?
- ¿Qué viajes parten desde el domicilio?
- ¿Cuántos pasajeros se trasladarán a un destino determinado, utilizando vehículo privado?

Por otra parte, también se tratará de analizar el efecto que tienen sobre las decisiones de los individuos otras variables diferentes de los atributos típicos de nivel de servicio como el tiempo de viaje, el coste y la frecuencia.

Este documento se divide en 3 capítulos. En el capítulo 1, detallaremos la base y los términos para entender qué es la demanda de transporte de pasajeros,

como se caracterizan los problemas del transporte y cuáles son las cuestiones importantes para seleccionar y desarrollar un modelo u otro.

En el capítulo 2 trataremos sobre el modelo clásico de transporte, las fases en las que se divide el desarrollo del modelo, y cuál de ellas tienen mayor importancia, centrándonos en la modelización de la distribución de viajes, que nos permitirá desarrollar un modelo sintético para usar en los datos obtenidos.

Por último, en el capítulo 3 se utilizará el **Modelo General de Gravedad** para estimar el número total de pasajeros que partirán de una parada de transporte público, a otra de destino. Todo ello se llevará a cabo a partir de una base de datos provista por la entidad de transporte público de la isla de Tenerife **TITSA**, apoyándonos en el uso de la herramienta QGIS para el análisis espacial de los datos.

Por último se presentan las principales conclusiones de este trabajo y aquellos aspectos que se consideran que deben ser objeto de futuras investigaciones.

## Planificación y modelización del transporte

La llegada de la informática de bajo coste y alta velocidad ha permitido que prácticamente se haya eliminado la potencia de cálculo como cuello de botella en la modelización del transporte. Las principales limitaciones son ahora humanas y técnicas: la planificación contemporánea del transporte, que requiere profesionales cualificados y experimentados, además de técnicas de modelización teóricamente sólidas con implementaciones competentes en software.

Los países emergentes son cada vez más importantes en la escena mundial, pero también sufren graves problemas de transporte. Ya no se trata sólo de la falta de carreteras para conectar las zonas rurales distantes con los mercados. De hecho, los nuevos problemas de transporte tienen algunas similitudes con los que prevalecen en el mundo post-industrializado: congestión, contaminación, etc. Sin embargo, tienen una serie de características muy distintivas que merecen un tratamiento específico: rentas relativamente bajas, urbanización, gran demanda de transporte público, escasez de recursos, incluidos el capital; los datos fiables y el personal cualificado.

Por todo ello, la previsión de la demanda es un elemento esencial en el análisis de los sistemas de transporte, el cuál:

- Se ocupa del comportamiento de los consumidores de servicios e instalaciones de transporte. Entre los usuarios se encuentran los viajeros y los expedidores de mercancías en los mercados de transporte urbano, interurbano e internacional.
- Analiza la respuesta de los usuarios a los cambios provocados por los nuevos servicios, las inversiones en infraestructuras y los cambios en las políticas de explotación y tarificación.

Los primeros estudios sobre las grandes inversiones en las redes regionales de carreteras y el rápido progreso de la tecnología computacional han llevado al desarrollo de procedimientos para predecir los viajes por orígenes y destinos y

los volúmenes de tráfico en los enlaces de una red.

Estos procedimientos de modelización se han ampliado y mejorado continuamente y se aplican en todo el mundo. Están bien documentados en la literatura (por ejemplo, [3], [4], [5]) y se apoyan en sistemas de software informático como el Sistema de Planificación del Transporte Urbano (UTPS) del Departamento de Transporte de EE.UU.

## 1.1. Características de los problemas de transporte

Los problemas de transporte se han extendido y agravado más que nunca tanto en los países industrializados como en los países en desarrollo. La escasez de combustible no es (temporalmente) un problema, pero el aumento general del tráfico por carretera y de la demanda de transporte ha provocado:

- Congestionamientos
- Retrasos
- Accidentes y problemas medioambientales mucho más allá de lo que se consideraba aceptable hasta ahora

Estos problemas no se han limitado a las carreteras y al tráfico de automóviles. El crecimiento económico parece haber generado niveles de demanda que superan la capacidad de la mayoría de las instalaciones de transporte.

Los largos periodos de falta de inversión en algunos modos de transporte y regiones han dado lugar a sistemas de suministro frágiles que parecen romperse cada vez que algo difiere ligeramente de las condiciones medias. No es probable que estos problemas desaparezcan en un futuro próximo.

Ha pasado suficiente tiempo con una mala o nula planificación del transporte para que sea necesario un gran esfuerzo de mejora de la mayoría de las formas de transporte, en contextos urbanos e interurbanos. Dado que los recursos no son ilimitados, este esfuerzo se beneficiará de decisiones cuidadosas y meditadas, orientadas a maximizar las ventajas de la nueva oferta de transporte al tiempo que se minimizan sus costes monetarios y sus efectos secundarios indeseables.

### 1.1.1. Demanda de transporte

La demanda de transporte viene derivada de una necesidad de la población en diferentes ámbitos, no es un fin en sí mismo. Con la posible excepción de las visitas turísticas, las personas se desplazan para satisfacer una necesidad

(trabajo, ocio, salud) realizando una actividad en determinados lugares. Esto es igualmente significativo para los movimientos de mercancías.

Para entender la demanda de transporte, hay que comprender la forma en que estas actividades se distribuyen en el espacio, tanto en el contexto urbano como en el regional.

1. Un buen sistema de transporte amplía las oportunidades de satisfacer estas necesidades.
2. Un sistema muy congestionado o mal conectado restringe las opciones y limita el desarrollo económico y social.

La demanda de servicios de transporte es muy cualitativa y diferenciada. Hay toda una serie de demandas específicas de transporte que se diferencian por la hora del día, el día de la semana, el propósito del viaje, el tipo de carga (ya sean personas, materiales, alimentos, etc.), la importancia de la velocidad y la frecuencia, etc. Un servicio de transporte sin los atributos correspondientes a esta demanda diferenciada puede resultar inútil.

Según [1], la demanda de transporte se desarrolla en el espacio. Parece una afirmación trivial, pero es la distribución de las actividades en el espacio lo que constituye la demanda de transporte. El enfoque más habitual para tratar el espacio consiste en dividir las áreas de estudio en zonas y codificarlas, junto con las redes de transporte, de forma adecuada para su tratamiento con ayuda de programas informáticos.

La espacialidad de la demanda suele provocar problemas de descoordinación que pueden afectar en gran medida al equilibrio entre la oferta y la demanda de transporte. Por otra parte, la concentración de la población y de la actividad económica en corredores bien definidos puede llevar a justificar económicamente un sistema de transporte masivo de alta calidad que no sería viable en una zona más dispersa.

Por último, es posible que un sistema de transporte pueda hacer frente a la demanda media de desplazamientos en una zona, pero que se rompa durante los periodos punta. Existen varias técnicas para intentar repartir los picos y promediar la carga del sistema:

- Horarios de trabajo flexibles.
- Escalonamiento de los horarios de trabajo.
- Tarificación adicional, etc.

Sin embargo, las variaciones de la demanda en las horas punta y fuera de ellas siguen siendo un problema central, y fascinante, en la modelización y planificación del transporte.

### 1.1.2. Oferta de transporte

La primera característica distintiva del suministro de transporte es que es un servicio y no un producto. Por lo tanto, no es posible almacenarlo, por ejemplo, para utilizarlo en momentos de mayor demanda. Un servicio de transporte debe consumirse en el momento y en el lugar en que se produce, pues de lo contrario se pierde su beneficio. Por eso es muy importante estimar la demanda con la mayor precisión posible para ahorrar recursos, adaptando la oferta de servicios de transporte a ella.

En términos muy generales, un sistema de transporte requiere una serie de activos fijos: la infraestructura; y una serie de unidades móviles: los vehículos. Es la combinación de éstos, junto con un conjunto de normas para su funcionamiento, lo que hace posible el movimiento de personas y mercancías.

La provisión de infraestructuras de transporte es especialmente importante desde el punto de vista de la oferta, ya que son "abultadas", es decir, no se puede proporcionar media pista de aterrizaje o un tercio de una estación de ferrocarril. En ciertos casos, puede haber margen para proporcionar un aumento gradual de la infraestructura para satisfacer la creciente demanda. De este modo, la dotación de infraestructuras puede ajustarse a la demanda y evitar inversiones tempranas innecesarias en instalaciones costosas. Esto es más difícil en otros ámbitos, como los aeropuertos, las líneas de metro, etc.

Un elemento adicional de distorsión es el número de efectos concomitantes o secundarios asociados a la producción de servicios de transporte: accidentes, contaminación y degradación del medio ambiente en general.

Una de las características más importantes de la oferta de transporte es la congestión. Se trata de un término difícil de definir, ya que todos creemos saber exactamente lo que significa. La congestión surge cuando los niveles de demanda se acercan a la capacidad de una instalación y el tiempo necesario para utilizarla (viajar por ella) aumenta muy por encima de la media en condiciones de baja demanda. En el caso de las infraestructuras de transporte, la inclusión de un vehículo adicional genera también un retraso suplementario para todos los demás usuarios. Este es el efecto externo de la congestión, percibido por los demás pero no por el conductor que la origina. Se trata de un coste que sistemas como la tarificación vial electrónica intentan internalizar para ayudar a una toma de decisiones más razonada por parte del individuo.

### 1.1.3. Planificación del transporte

En términos generales, la función de la planificación de los transportes es garantizar la satisfacción de una determinada demanda  $\mathbf{D}$  de movimientos de personas y mercancías con diferentes propósitos de viaje, en diferentes momentos del día y del año, utilizando varios modos, dado un sistema de transporte con una determinada capacidad operativa. El sistema de transporte puede considerarse compuesto por:

- Una **infraestructura** (por ejemplo, una red de carreteras).
- Un **sistema de gestión** (es decir, un conjunto de normas, por ejemplo la conducción por la derecha, y estrategias de control, por ejemplo, en los semáforos).
- Un conjunto de **modos** de transporte y sus operadores. El modo hace referencia al tipo de transporte a utilizar, ya sea vehículo público o privado, bicicleta, caminando, etc. Este término se utilizará en muchas ocasiones a lo largo de este capítulo y los siguientes.

## 1.2. Elección de enfoques de modelización

La aceptabilidad de la modelización, o de un determinado enfoque de modelización, dentro de un estilo de decisión es muy importante. Los modelos que acaban siendo ignorados por los responsables de la toma de decisiones no sólo representan un desperdicio de recursos y esfuerzos, sino que dan lugar a analistas y planificadores frustrados. Se proponen varias características de los problemas y modelos de transporte que deben tenerse en cuenta a la hora de especificar un enfoque analítico:

1. **Se requiere precisión y exactitud.**
2. **El contexto de la toma de decisiones**
3. **Nivel de detalle requerido**
4. **La disponibilidad de datos adecuados**, su estabilidad y las dificultades para prever sus valores futuros.
5. **El estado de la técnica en materia de modelización** para un determinado tipo de intervención en el sistema de transporte. Este a su vez puede subdividirse en:
  - Riqueza de comportamiento.
  - Trazabilidad matemática e informática.
  - Disponibilidad de buenos algoritmos de solución.
6. **Recursos disponibles para el estudio.**
7. **Requisitos de procesamiento de datos**
8. **Niveles de formación y competencias de los analistas.**
9. **Perspectiva y alcance de la modelización**

### 1.3. Cuestiones de la modelización del transporte

Ya hemos identificado las interacciones entre los problemas de transporte, los estilos de decisión y los enfoques de modelización. Ahora tenemos que discutir algunas de las cuestiones críticas de la modelización que son relevantes para la elección del modelo. Estas cuestiones abarcan algunos puntos generales como el papel de la teoría y los datos, la especificación del modelo y la calibración. Pero quizá las opciones más importantes sean las que se refieren a la utilización de enfoques agregados o desagregados, modelos de sección transversal o de series temporales, y técnicas de preferencias reveladas o declaradas. Éstas últimas no estarán dentro del marco teórico y de estudio del trabajo.

#### 1.3.1. Especificaciones del modelo

Podemos destacar tres especificaciones fundamentales:

**Estructura del modelo.** Los modelos desagregados suelen tener parámetros que representan aspectos de la estructura del modelo, y las ampliaciones de la metodología logradas a mediados de los años 80 han permitido estimar formas de modelos cada vez más generales. Sin embargo, como se ha señalado en [14], aunque se podría suponer que, en última instancia, todas las cuestiones relativas a la forma del modelo podrían resolverse mediante pruebas empíricas, dicha resolución no es posible ni adecuada.

**Forma funcional.** ¿Es posible utilizar formas lineales o el problema requiere postular funciones no lineales más complejas? Estas últimas pueden representar el sistema de interés con mayor precisión, pero ciertamente serán más exigentes en términos de recursos y técnicas para la calibración y uso del modelo.

**Especificación de las variables.** Este es el significado más habitual de la cuestión de la especificación: qué variables utilizar y cómo (de qué forma) deben entrar en un modelo determinado. Los métodos para avanzar en esta cuestión van desde el uso deductivo (“constructivo”) de la teoría hasta el análisis estadístico inductivo de los datos mediante transformaciones.

#### 1.3.2. Calibración, validación y uso del modelo

Un modelo puede representarse simplemente como una función matemática de variables  $X$  y parámetros  $\theta$ , como:

$$Y = f(X, \theta)$$

Es interesante mencionar que los conceptos gemelos de **calibración** y **estimación** de modelos han tomado tradicionalmente un significado diferente en el campo del transporte.

La **calibración** de un modelo requiere la elección de sus parámetros, de los que se supone que tienen un valor no nulo, con el fin de optimizar una o varias medidas de bondad de ajuste, que son función de los datos observados.

La **estimación** consiste en encontrar los valores de los parámetros que hacen que los datos observados sean más probables según la especificación del modelo; en este caso, uno o varios parámetros pueden considerarse no significativos y dejarse fuera del modelo. La estimación también considera la posibilidad de examinar empíricamente ciertas cuestiones de especificación.

A la hora de validar un modelo, se requiere comparar las predicciones del modelo con información no utilizada durante el proceso de estimación del modelo. Una de las primeras tareas a las que se enfrenta un modelizador es decidir qué variables va a predecir el modelo y cuáles son posiblemente necesarias como entradas para el mismo. Esto implica inmediatamente un cierto grado de error e incertidumbre que, por supuesto, se ve agravado por otros errores que también son inherentes a la modelización. Así, el principal uso de los modelos en la práctica es la previsión condicional: el modelo producirá estimaciones de las variables dependientes dado un conjunto de variables independientes. De hecho, las previsiones típicas son condicionales de dos maneras (véase [6]):

- En relación con los valores asignados a las variables políticas del plan, cuyo impacto se está probando con el modelo.
- En relación con los valores asumidos de otras variables.



## La estructura del modelo clásico de transporte

Años de experimentación y desarrollo han dado lugar a una estructura general que se ha denominado modelo de transporte clásico. Esta estructura es, en efecto, un resultado de la práctica de los años 60, pero ha permanecido más o menos inalterada a pesar de las importantes mejoras en las técnicas de modelización desde entonces.

La forma general del modelo se representa en la figura (2.1). El planteamiento parte de la consideración de un sistema de zonificación y red, y de la recogida y codificación de datos de planificación, calibración y validación. Estos datos incluirían los niveles del año base para la población de diferentes tipos en cada zona del área de estudio, así como los niveles de actividad económica, incluyendo el empleo, el espacio comercial y las instalaciones educativas y recreativas. Estos datos se utilizan para estimar un modelo del número total de viajes generados y atraídos por cada zona del área de estudio (generación de viajes). El siguiente paso es la asignación de estos viajes a destinos concretos, es decir, su distribución en el espacio, produciendo así una matriz de viajes.

La siguiente etapa suele consistir en la modelización de la elección del modo de transporte, lo que da lugar al reparto modal, es decir, a la asignación de los viajes de la matriz a los distintos modos. Finalmente, la última etapa del modelo clásico requiere la asignación de los viajes de cada modo a sus redes correspondientes: normalmente el transporte privado y el público.

El modelo clásico (2.1) se presenta como una secuencia de cuatro submodelos: generación, distribución, reparto modal y asignación. En general, se reconoce que las decisiones de viaje no se toman realmente en este tipo de secuencia; una opinión contemporánea es que la “ubicación” de cada submodelo depende de la forma de la función de utilidad que se supone que rige todas estas elecciones de viaje. Además, se considera que el modelo de cuatro etapas concentra la atención en una gama limitada de respuestas de los viajeros. El pensamiento actual

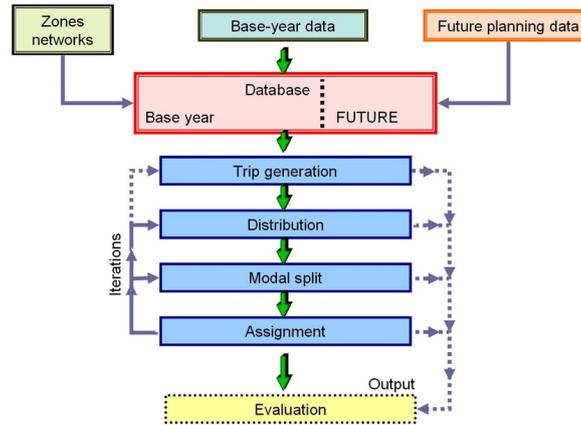


Figura 2.1: El modelo clásico de transporte en cuatro etapas.

requiere un análisis de una gama más amplia de respuestas a los problemas y planes de transporte. Por ejemplo, ante un aumento de la congestión, un viajero puede responder con una serie de cambios sencillos en:

- La **ruta** seguida para evitar la congestión o aprovechar los nuevos enlaces; esto incluye la elección del lugar de aparcamiento o la combinación de servicios en el caso del transporte público; .
- El **modo** de transporte utilizado para llegar al destino.
- La **hora** de salida para evitar la zona más congestionada de la hora punta.
- El **destino** del viaje a una zona menos congestionada.
- La **frecuencia** de los desplazamientos hacia los diversos destinos existentes.

Además, a más largo plazo se producen otras respuestas más complejas, como los cambios en el empleo, la ubicación residencial, la elección de las zonas comerciales, etc; todas ellas responderán, al menos parcialmente, a los cambios en la accesibilidad proporcionada por el sistema de transporte. A pesar de estos comentarios, el modelo secuencial de cuatro etapas proporciona un punto de referencia para contrastar métodos alternativos.

Una vez preparados los escenarios y planes realistas para las pruebas, se vuelve a ejecutar la misma secuencia de modelos para simular su rendimiento. A continuación, se comparan los costes y beneficios, independientemente de cómo se midan, de los distintos planes en los diferentes escenarios; la idea es elegir el programa más atractivo de inversiones y políticas de transporte que satisfaga la demanda de movimiento en la zona de estudio.

Una cuestión importante en el modelo clásico de cuatro etapas es el uso coherente de las variables que afectan a la demanda. Además, la aplicación posterior del modelo de asignación puede dar lugar a un nuevo conjunto de tiempos

de viaje; se verá que, en general, la retroalimentación ingenua del modelo no conduce a un conjunto estable de modelos de distribución, reparto modal y asignación con tiempos de viaje coherentes. A continuación, se introducirán los cuatro submodelos que caracterizan al modelo clásico de transporte, definiéndolos y desglosando sus respectivas características.

## 2.1. Modelización de la generación de viajes

Como se vió anteriormente, la etapa de generación de viajes del modelo de transporte clásico tiene como objetivo predecir el número total de viajes generados por ( $O_i$ ) y dirigidos a cada zona ( $D_j$ ) del área de estudio. Esto se puede conseguir de varias maneras: partiendo de los viajes de los individuos u hogares que residen en cada zona o directamente de algunas de las propiedades de las regiones o áreas de estudio: población, empleo, número de coches, etc.

El tema también se ha considerado como un problema de elección de la frecuencia de los viajes: *¿cuántos viajes de compras (o de otro tipo) realizará este tipo de persona durante una semana representativa?* Para ello se suelen utilizar modelos de elección discreta, y se plantea en términos como: *¿cuál es la probabilidad de que este tipo de persona realice cero, uno, dos o más viajes con este propósito a la semana?*

En este capítulo nos centramos en el primer enfoque (es decir, la predicción de los totales  $O_i$  y  $D_j$  a partir de datos sobre atributos socioeconómicos de los hogares). Comenzaremos por definir algunos conceptos básicos y procederemos a examinar algunos de los factores que afectan a la generación y atracción de viajes.

### 2.1.1. Definiciones básicas

Entender los términos básicos que se utilizan en el ámbito del transporte dará una mayor noción para contestar preguntas como:

- ¿Cuál es el acontecimiento básico de interés para nuestros modelos, las actividades que implican un breve periodo de estancia en un lugar, el desplazamiento de un lugar a otro?, o,
- ¿una secuencia de estos viajes y estancias que empiezan y terminan en casa, constituyen una excursión o un viaje?
- ¿Qué viajes no tienen como base el domicilio?

El modelista debe tener en cuenta todas estas opciones a la hora de elegir de la siguiente lista para desarrollar y aplicar un conjunto de modelos. Según [1], tenemos:

- **Viaje o travesía:** Se trata de un movimiento unidireccional desde un punto de origen hasta un punto de destino. Normalmente interesan todos los desplazamientos en vehículo. A menudo se ignoran los desplazamientos a pie inferiores a un determinado umbral definido por el estudio (por ejemplo, 300 metros o tres manzanas), así como los desplazamientos de los niños menores de cinco años de edad. Sin embargo, este énfasis está cambiando, ya que se presta más atención a los viajes no motorizados y debido a los requisitos del marco de actividad recomendado para las encuestas de movilidad.
- **Viaje basado en el hogar (HB):** Se trata de un viaje en el que el hogar del viajero es el origen o el destino del viaje. Obsérvese que para los visitantes procedentes de otra ciudad, su hospedaje actúa como hogar temporal en la mayoría de los estudios.
- **Viaje no basado en el hogar (NHB):** Por el contrario, en este caso ninguno de los dos extremos del viaje es el hogar del viajero.
- **Producción del viaje:** Es el final del viaje basado en el hogar y es el origen del viaje no basado en el hogar.

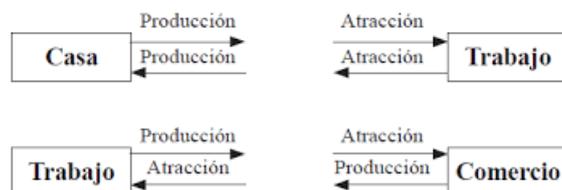


Figura 2.2: Producción y atracción de viajes

- **Atracción de viaje:** Es el extremo no doméstico del viaje basado en el hogar y es el destino de un viaje no basado en el hogar (véase la figura (2.2)).
- **Generación de viajes:** Suele definirse como el número total de viajes generados por los hogares de una zona, ya sea **HB** o **NHB**. Esto es lo que producen la mayoría de los modelos y la tarea consiste en asignar los viajes **NHB** a otras zonas como producción de viajes.
- **Residencia temporal:** Periodo corto de estancia en un lugar determinado. Suele tener una finalidad asociada a esta estancia: trabajo, estudio, compras, ocio, etc.
- **Actividad:** Un esfuerzo o interés a menudo asociado a un propósito como el anterior, pero no necesariamente vinculado a un lugar fijo. Se puede elegir ir de compras o al cine en diferentes lugares.
- **Tour o cadena de viajes:** Un conjunto de estancias y viajes enlazados. Los tres últimos conceptos se corresponden mejor con la idea de viaje como demanda derivada (es decir, que depende en gran medida de la demanda

de otras actividades), pero inicialmente fueron utilizados principalmente por los modelizadores de elección discreta en la práctica (véase [7]). Los modelos contemporáneos, sobre todo los del tipo de frecuencia de viajes, se interesan más por los desplazamientos.

### 2.1.2. Caracterización de los viajes

La caracterización de los viajes se puede llevar a cabo:

- I) **Por objetivos.** En la práctica, se ha comprobado que se puede obtener una mejor comprensión de los modelos de generación de viajes y desplazamientos si se identifican y modelizan por separado los viajes con diferentes propósitos. En el caso de los viajes de **HB**, se han empleado varias categorías:
- Viajes al trabajo.
  - Viajes a la escuela o al colegio (viajes de educación).
  - Viajes de compras.
  - Viajes sociales y recreativos.
  - Viajes de acompañamiento (para acompañar o recoger a otra persona).
  - Otros viajes.

Los dos primeros suelen denominarse viajes obligatorios y todos los demás se denominan viajes discrecionales (u opcionales). Esta última categoría engloba todos los viajes realizados con fines menos rutinarios, como la salud y los asuntos personales (necesidad de obtener un pasaporte o un certificado). Tener en cuenta que los contextos sociales y culturales pueden modificar la importancia de los distintos tipos de viajes y, por tanto, la clasificación más adecuada. Los viajes **NHB** se separan a veces en “por negocios” y “otros”, pero a menudo se mantienen como una categoría única porque sólo suponen el 15-20 % del total de los viajes.

- II) **Por el momento del día.** Los viajes se clasifican a veces en viajes en horas punta y fuera de horas punta; la proporción de desplazamientos por diferentes motivos suele variar mucho según la hora del día. Este tipo de clasificación, aunque importante, se complica cuando interesan los recorridos en lugar de los viajes, ya que un recorrido completo puede comprender viajes realizados en varios momentos del día.
- III) **Por el tipo de persona.** Esta es otra clasificación importante, ya que el comportamiento individual de los viajes depende en gran medida de los atributos socioeconómicos. Se suelen emplear las siguientes categorías:
- Nivel de ingresos (por ejemplo, tres estratos: ingresos bajos, medios y altos).
  - Propiedad de vehículos (normalmente tres estratos: 0, 1 y 2 o más coches).
  - Tamaño y estructura del hogar (por ejemplo, seis estratos en los estudios británicos clásicos).

Es importante tener en cuenta que el número total de estratos puede aumentar muy rápidamente y esto puede tener fuertes implicaciones en términos de requisitos de datos, calibración del modelo y uso. Por este motivo, suele ser necesario realizar compensaciones, ajustes y agregaciones (véase la discusión en [15]).

### 2.1.3. Factores que afectan a la generación de viajes

En la modelización de la generación de viajes se suele estar interesados no sólo en los viajes de personas sino también en los de mercancías. Por esta razón, se han requerido modelos para cuatro grupos principales (es decir, viajes personales y de carga, producciones de viajes y atracciones) que se han necesitado normalmente. A continuación, se examinarán brevemente algunos factores que se han considerado importantes en los estudios prácticos.

No obstante, no se tratará la modelización de la generación de viajes de mercancías. Presentaremos un esquema para facilitar la visualización de los factores:

– **Producción de viajes por persona**

- a) Nivel de ingreso
- b) Tamaño de la familia
- c) Accesibilidad
- d) Propiedad vehicular
- e) Densidad residencial

– **Atracción de viajes de personas**

- a) Área disponible para servicios industriales y comerciales
- b) Número de empleos públicos
- c) Número de establecimientos educativos

– **Viajes de ida y vuelta en producción y atracción**

- a) Número de empleados
- b) Número de ventas
- c) Superficie cubierta de la empresa
- d) Superficie total de la empresa

## 2.2. Modelización de la distribución de viajes: *Modelo de Gravedad*

Se ha visto cómo los modelos de generación de viajes pueden utilizarse para estimar el número total de viajes que emanan de una zona (orígenes, producciones) y los que se atraen a cada zona (destinos, atracciones). Las producciones y

atracciones proporcionan una idea del nivel de realización de viajes en un área de estudio, pero esta rara vez es suficiente para la modelización y la toma de decisiones. Lo que se necesita es una mejor idea del patrón de realización de viajes, desde dónde se realizan los viajes, los modos de transporte elegidos y las rutas que se toman.

El patrón de viajes puede representarse, en esta fase, al menos de dos maneras diferentes.

- La primera es una “matriz de viajes” o “tabla de viajes”.
- La segunda forma de presentar un patrón de viaje es considerar los factores que generan y atraen los viajes, es decir, sobre una base de Producción-Atracción (**P-A**), tratándose generalmente el hogar como el extremo “productor”, y el trabajo, la tienda, etc, como el extremo “atrayente”.

La distribución de viajes suele considerarse un problema agregado con un modelo agregado para su solución. Sin embargo, la elección del destino también puede tratarse como un problema de elección discreta (desagregada), y tratarse con modelos a nivel del individuo.

### 2.2.1. Definiciones y notación

Actualmente es habitual representar el patrón de viajes en un área de estudio mediante una matriz de viajes. Se trata esencialmente de una matriz bidimensional de celdas en la que las filas y las columnas representan cada una de las  $z$  zonas del área de estudio (incluidas las zonas externas), como se muestra en la tabla (2.1).

Las celdas de cada fila  $i$  contienen los viajes con origen en esa zona, que tienen como destino las zonas de las columnas correspondientes. La diagonal principal corresponde a los viajes intrazonales. Por lo tanto  $T_{ij}$  es el número de viajes entre el origen  $i$  y el destino  $j$ ; la matriz total es  $\{T_{ij}\}$  o **T**;  $O_i$  es el número total de viajes originados en la zona  $i$ , y  $D_j$  es el número total de viajes destinados a la zona  $j$ .  $P_i$  es el número de viajes producidos o generados en una zona  $i$  y  $A_j$  los atraídos a la zona  $j$ .

Se utilizarán letras minúsculas,  $t_{ij}$ ,  $o_i$  y  $d_j$  para indicar las observaciones de una muestra o de un estudio anterior; las letras mayúsculas representarán nuestro objetivo, o los valores que intentamos modelar para el periodo de modelización correspondiente.

Las matrices pueden desagregarse aún más, por ejemplo, por tipo de persona ( $n$ ) y/o por modo de transporte ( $k$ ). Por lo tanto:

- $T_{ij}^{kn}$  son los viajes de  $i$  a  $j$  en el modo de transporte  $k$  y el tipo de persona  $n$ ;
- $O_i^{kn}$  es el número total de viajes originados en la zona  $i$  por el modo de transporte  $k$  y el tipo de persona  $n$ , y así sucesivamente.

Destinos	1	2	3	...j	...z	$\sum_i T_{ij}$
Orígenes	1	2	3	...j	...z	$\sum_i T_{ij}$
1	$T_{11}$	$T_{12}$	$T_{13}$	$\dots T_{1j}$	$\dots T_{1z}$	$O_1$
2	$T_{21}$	$T_{22}$	$T_{23}$	$\dots T_{2j}$	$\dots T_{2z}$	$O_2$
3	$T_{31}$	$T_{32}$	$T_{33}$	$\dots T_{3j}$	$\dots T_{3z}$	$O_3$
			$\vdots$			
I	$T_{i1}$	$T_{i2}$	$T_{i3}$	$\dots T_{ij}$	$\dots T_{iz}$	$O_i$
			$\vdots$			
Z	$T_{z1}$	$T_{z2}$	$T_{z3}$	$\dots T_{zj}$	$\dots T_{zz}$	$O_z$
$\sum_i T_{ij}$	$D_1$	$D_2$	$D_3$	$\dots D_j$	$\dots D_z$	$\sum_{ij} T_{ij} = T$

Tabla 2.1: Una forma general de una matriz de viaje bidimensional

La suma sobre subíndices o superíndices se indicará implícitamente por omisión, por ejemplo:

$$T_{ij}^n = \sum_k T_{ij}^{kn} \quad , \quad T = \sum_{ij} T_{ij} \quad \text{y} \quad t = \sum_{ij} t_{ij}$$

En algunos casos puede ser interesante distinguir la proporción de viajes que utilizan un modo concreto y el coste del viaje entre dos puntos:

- $p_{ij}^k$  es la proporción de viajes de  $i$  a  $j$  en el modo de transporte  $k$ .
- $c_{ij}^k$  es el coste de viaje de  $i$  a  $j$  en el modo de transporte  $k$ .

Además se deben cumplir las siguientes condiciones:

$$\sum_j T_{ij} = O_i \tag{2.1}$$

$$\sum_i T_{ij} = D_j \tag{2.2}$$

Si se dispone de información fiable para estimar tanto  $O_i$  como  $D_j$ , el modelo debe satisfacer ambas condiciones; en este caso se dice que el modelo está doblemente restringido. En algunos casos, sólo se dispondrá de información sobre una de estas restricciones y por tanto se dirá que el modelo está restringido

individualmente.

Por otro lado, el coste puede considerarse en términos de distancia, tiempo o unidades de dinero. A menudo es conveniente utilizar una medida que combine todos los atributos principales relacionados con la desutilidad de un viaje y que normalmente se denomina coste generalizado del viaje. Suele ser una función lineal de los atributos del viaje ponderada por coeficientes que intentan representar su importancia relativa según la percepción del viajero. Una posible representación para el modo de transporte  $k$  es (omitiendo el superíndice  $k$  para simplificar):

$$C_{ij} = a_1 t_{ij}^v + a_2 t_{ij}^w + a_3 t_{ij}^t + a_4 t_{ij}^n + a_5 F_{ij} + a_6 \phi_j + \delta \quad (2.3)$$

donde

- $t_{ij}^v$  es el tiempo de viaje, a través de un vehículo, entre  $i$  y  $j$ ;
- $t_{ij}^w$  es el tiempo que se tarda en llegar desde  $i$  a  $j$  partiendo desde una parada de guaguas (estaciones) o un área de aparcamientos;
- $t_{ij}^t$  es el tiempo de espera en las paradas (o el tiempo de búsqueda de aparcamiento);
- $t_{ij}^n$  es el tiempo de intercambio (pasar de un modo de transporte a otro a lo largo del viaje), si lo hay;
- $F_{ij}$  es una carga monetaria: la tarifa cobrada por viajar entre  $i$  y  $j$  o el coste de usar el coche para ese viaje, incluyendo cualquier peaje o tasa de congestión;
- $\phi_j$  es un coste de terminal (normalmente de aparcamiento) asociado al viaje de  $i$  a  $j$ ;
- $\delta$  es una penalización modal, un parámetro que representa todos los demás atributos no incluidos hasta ahora en la medida generalizada como, por ejemplo, la seguridad, el confort y la comodidad;
- $a_1, \dots, a_6$  son los pesos asignados a cada elemento del coste; tienen dimensiones apropiadas para la conversión de todos los atributos a unidades comunes, por ejemplo, dinero o tiempo.

En definitiva, un modelo de distribución trata de estimar el número de viajes en cada una de las celdas de la matriz a partir de cualquier información disponible. En este caso, según con los datos que se trabajará en el capítulo 3, se utilizará uno de los modelos pertenecientes a la *familia de modelos de gravedad*.

### 2.2.2. Generación del Modelo de Gravedad

Antes de continuar, considérese un sistema compuesto por un gran número de elementos distintos. Una descripción completa de tal sistema requiere la especificación completa de sus **microestados** (variables explicativas de cada individuo), ya que cada uno es distinto y separable. Esto implicaría, por ejemplo,

- La identificación de cada viajero individual
- Su origen y destino
- Modo de transporte a utilizar (vehículo público o privado, bicicleta, caminando)
- Tiempo de viaje, etc.

Sin embargo, para muchos fines prácticos puede ser suficiente trabajar sobre la base de una especificación más agregada o **mesoestado**; siguiendo el ejemplo, un **mesoestado** puede especificar simplemente el número de viajes entre cada origen y cada destino. En general, habrá numerosos y diferentes microestados que produzcan el mismo mesoestado.

Siempre existe un nivel de agregación aún mayor, un **macroestado**, por ejemplo el número total de viajes en determinados enlaces, o el total de viajes generados y atraídos a cada zona. Para obtener medidas fiables de la actividad de realización de viajes suele ser más fácil realizar observaciones a este nivel superior de agregación. De hecho, la mayor parte de nuestra información actual sobre un sistema se encuentra precisamente en este nivel.

Es posible demostrar (véase [8]) que el número de microestados  $W\{T_{ij}\}$  asociados al meso estado  $T_{ij}$  viene dado por:

$$W\{T_{ij}\} = \frac{T!}{\prod_{ij} T_{ij}!} \quad (2.4)$$

Como se supone que todos los microestados son igualmente probables, el mesoestado más probable sería el que puede generarse de un mayor número de maneras. Por lo tanto, lo que se necesita es una técnica para identificar los valores de  $\{T_{ij}\}$  que maximizan  $W$  en (2.4). Por comodidad buscamos maximizar una función monótona de  $W$ , concretamente  $\log(W)$ , ya que ambos problemas tienen el mismo máximo. Por lo tanto:

$$\log W = \log \left( \frac{T!}{\prod_{ij} T_{ij}!} \right) = \log(T!) - \sum_{ij} \log(T_{ij}!) \quad (2.5)$$

La aproximación (corta) de Stirling para  $\log(X!) = X \log(X) - X$ , puede utilizarse para facilitar la optimización de:

$$\log(W) = \log(T!) - \sum_{ij} (T_{ij} \log(T_{ij}) - T_{ij}) \quad (2.6)$$

Normalmente, el término  $\log(T!)$  es una constante, por lo que puede omitirse en el problema de optimización. El resto de la ecuación suele denominarse función de la entropía (realizando la primera derivada sobre  $W$ ):

$$\log(W') = - \sum_{ij} (T_{ij} \log(T_{ij}) - T_{ij}) \quad (2.7)$$

La maximización del  $\log(W')$ , sujeta a las restricciones correspondientes a nuestro conocimiento sobre los **macroestados**, nos permite generar modelos para estimar los **mesoestados** más probables, en nuestro caso la matriz  $\mathbf{T}$  más probable. La clave de este método de generación de modelos es, por tanto, la identificación de las descripciones de micro, meso y macroestado adecuadas, junto con las restricciones de nivel macro que debe cumplir la solución del problema de optimización.

En algunos casos, puede haber información adicional en forma de valores previos o antiguos para los **mesoestados**, por ejemplo, una matriz de viaje  $\mathbf{t}$  no actualizada. El problema puede refundirse con esta información y la función objetivo revisada sería:

$$\log(W'') = - \sum_{ij} \left( T_{ij} \log \left( \frac{T_{ij}}{t_{ij}} \right) - T_{ij} + t_{ij} \right) \quad (2.8)$$

Se trata de una función interesante en la que cada elemento del sumatorio toma el valor cero si  $T_{ij} = t_{ij}$  y, en caso contrario, es un valor positivo que aumenta con la diferencia entre  $\mathbf{T}$  y  $\mathbf{t}$ . El modelo resultante es el que tiene los *mesoestados* más cercanos a los *mesoestados* previos, en el sentido de la ecuación (2.8), y que satisface las restricciones de los *macroestados*.

Una vez consideradas las definiciones de micro, meso y macroestados descritas anteriormente, el problema se convierte en la maximización de  $\log(W')$ , sujeta a los siguientes dos conjuntos de restricciones correspondientes a los estados *meso*:

$$O_i - \sum_j T_{ij} = 0 \quad (2.9)$$

$$D_j - \sum_i T_{ij} = 0 \quad (2.10)$$

Estos dos conjuntos de restricciones reflejan nuestro conocimiento sobre la producción de viajes y las atracciones en las zonas del área de estudio. Sólo nos interesan las entradas de la matriz que pueden interpretarse como viajes, por lo que necesitamos introducir la restricción adicional de que:

$$T_{ij} \geq 0$$

El problema de maximización restringida se puede manejar formando el Lagrangiano:

$$L = \log(W') + \sum_i \alpha'_i \left( O_i - \sum_i T_{ij} \right) + \sum_j \alpha''_j \left( D_j - \sum_i T_{ij} \right) \quad (2.11)$$

Los multiplicadores de Lagrange  $\alpha_i$  y  $\alpha_j$  son las variables duales de las restricciones de generación y atracción de viajes y se relacionan con las variaciones de entropía para una variación unitaria en la generación y atracción de viajes. Tomando las primeras derivadas parciales con respecto a  $T_{ij}$  e igualándolas a cero obtenemos:

$$\frac{\partial L}{\partial T_{ij}} = -\log(T_{ij}) - \alpha'_i - \alpha''_j = 0 \quad (2.12)$$

por lo tanto,

$$T_{ij} = \exp(-\alpha'_i - \alpha''_j) = \exp(-\alpha'_i) \exp(-\alpha''_j) \quad (2.13)$$

Los valores de los multiplicadores de Lagrange son fáciles de encontrar, haciendo un simple cambio de variables:

$$A_i O_i = \exp(-\alpha'_i) \quad \text{y} \quad B_j D_j = \exp(-\alpha''_j)$$

Obtenemos así que:

$$T_{ij} = A_i O_i B_j D_j \quad (2.14)$$

Por otro lado, el uso de  $-\log(W'')$  como función objetivo genera el modelo:

$$T_{ij} = A_i O_i B_j O_j t_{ij} \quad (2.15)$$

que se trata del modelo básico de Furness<sup>1</sup>. La versión que da lugar a la ecuación (2.14) corresponde al caso en que no hay información previa, por ejemplo,

<sup>1</sup> Es un método que utiliza factores de crecimiento y considera que en el futuro los patrones de viaje se conservan sustancialmente, pero el volumen de viajes se incrementa de acuerdo al crecimiento de las zonas de producción y atracción de viajes

todo  $t_{ij} = 1$ . Estos dos modelos se acercan al modelo de gravedad, pero hay que precisarlos aun más. Lo que falta es el término de la función de disuasión ( $\exp(-\beta c_{ij})$ )<sup>2</sup>. Su introducción requiere una restricción adicional:

$$\sum_{ij} T_{ij} c_{ij} = C \quad (2.16)$$

donde  $C$  es el gasto total (desconocido) en viajes en el sistema (en unidades de coste generalizadas si se utilizan). Reformulando esta restricción

$$C - \sum_{ij} T_{ij} c_{ij} = 0 \quad (2.17)$$

se puede maximizar  $\log(W')$  sujeto a (2.9), (2.10) y (2.17), y utilizando la misma técnica de optimización restringida es posible obtener el Lagrangiano:

$$L = \log(W') + \sum_i \alpha'_i \left( O_i - \sum_j T_{ij} \right) + \sum_j \alpha''_j \left( D_j - \sum_i T_{ij} \right) + \beta \left( C - \sum_{ij} T_{ij} c_{ij} \right) \quad (2.18)$$

De nuevo, tomando sus primeras derivadas parciales con respecto a  $T_{ij}$  e igualándolas a cero se obtiene

$$\frac{\partial L}{\partial T_{ij}} = -\log(T_{ij}) - \alpha'_i - \alpha''_j - \beta c_{ij} = 0 \quad (2.19)$$

entonces,

$$T_{ij} = \exp(-\alpha'_i - \alpha''_j - \beta c_{ij}) = \exp(-\alpha'_i) \exp(-\alpha''_j) \exp(-\beta c_{ij}) \quad (2.20)$$

Haciendo el mismo cambio de variables que antes se obtiene:

$$T_{ij} = A_i O_i B_j D_j \exp(-\beta c_{ij}) \quad (2.21)$$

que es el modelo clásico de gravedad. Los valores de los factores de equilibrio pueden derivarse de las restricciones como:

$$A_i = \frac{1}{\sum_j B_j D_j \exp(-\beta c_{ij})} \quad \text{y} \quad B_j = \frac{1}{\sum_i A_i O_i \exp(-\beta c_{ij})}$$

El valor de  $\beta$  está relacionado con la satisfacción de la condición ((2.17)). En general,  $C$  sólo puede estimarse y, por tanto,  $\beta$  se deja como parámetro de calibración para ajustar el modelo a cada zona concreta.

<sup>2</sup> Esta función recibe a menudo el nombre de "función de disuasión" porque representa el desincentivo para viajar a medida que aumenta la distancia (el tiempo) o el coste.

Por tanto, los valores de  $\beta$  no pueden tomarse prestados fácilmente de un lugar a otro. Una primera estimación útil para el valor de  $\beta$  es una sobre el coste medio del viaje; en efecto,  $\beta$  se mide precisamente en unidades inversas del coste del viaje.

## Análisis del flujo de pasajeros en Tenerife

Una vez visto y estudiado los conceptos de los dos capítulos anteriores, se tratará a continuación de analizar los datos de la movilidad de pasajeros en la isla de Tenerife. Para ello, se trabajará con un conjunto de datos proporcionado por la entidad **TITSA**, perteneciente al Cabildo Insular de Tenerife que da servicio de transporte público en superficie (guaguas y tranvía) en la isla de Tenerife. La idea de éste capítulo será la de, primeramente, visualizar y describir la fluctuación de pasajeros entre todas las paradas existentes en la isla de Tenerife, para posteriormente construir una modelización que se ajuste a los datos proporcionados y permita estimar cuál será el número de pasajeros en determinadas paradas claves y sus respectivos destinos.

### 3.1. Fuente de datos

Los datos de la investigación se pueden dividir en dos conjuntos: información sobre las paradas de toda la isla de Tenerife, y un segundo conjunto de datos con la información de entrada y salida de una guagua o tranvía entre los días 25-30 de Diciembre y los días 1, 2 y 6 de Enero. El primer conjunto sobre la información de las paradas se desgloza en:

- *Parada*: Número de la parada de guaguas/tranvía.
- *Parada\_Name*: Nombre de la parada de guaguas/tranvía.
- *Latitud y Longitud*.
- *Radio*: Radio de la circunferencia, dada en metros, que delimita con otras paradas de la zona.
- *Ubicación*: Nombre de la calle, con su respectivo número; Autovía o Autopista.
- *CP*: Código postal de la zona geográfica a la que pertenece la parada de guaguas/tranvía.

- *Municipio*: División territorial administrativa a la que pertenece la parada de guaguas/tranvía.
- *Tipo\_Parada*: Según la zona geográfica, teniendo líneas Interurbanas (Zona Sur, Zona Norte, Otros), línea Aeroexpress, líneas Metropolitanas, líneas Urbanas, línea Tuwawa y otros servicios <sup>1</sup>; tenemos que el tipo de parada puede ser de:
  - TITSA Oficial
  - Transporte la Esperanza
  - Tranvía
  - Otro
- *Estado\_Parada*: Si la parada de guaguas/tranvía sigue estando activa, ha sido dada de baja u otro estado.

El segundo conjunto de datos sobre la información obtenida cuando un pasajero sube y baja del transporte se desgloza en:

- *Perfil\_Usuario*: Tipo de bono utilizado por el pasajero de entrada (pudiendo no utilizar ningún bono), que puede ser uno de los siguientes:
  - Desconocido
  - Discapacitado C, Discapacitado General
  - Empleado MTSA (Metrotenerife.Tranvía de Tenerife), Empleado TITSA
  - Estudiante Universitario
  - Familia Numerosa C, Familiar personal TITSA
  - General
  - Joven menor de 30, Joven menor de 30 (La Gomera)
  - Jubilado C, Jubilado TITSA
  - Mayor de 65, Niño menor de 10
  - Residente Canario, Residente Canario Desempleado
  - Trabajador Aeropuerto
- *Fecha\_Subida, Hora\_Subida, Parada\_Subida*
- *Fecha\_Bajada, Hora\_Bajada, Parada\_Bajada*
- *Num\_Transbordos*: Número de transbordos realizados por el pasajero, realizando como máximo 3 transbordos.

<sup>1</sup> *Líneas Interurbanas*: líneas que conectan las principales poblaciones de la isla. Estos servicios pueden ser convencionales, parando en todas las paradas (si se solicita) o express, sin paradas intermedias; *Línea Aeroexpress*: 3 líneas express:

- Línea 20 Santa Cruz - La Laguna - Aeropuerto Norte
- Línea 30 Puerto de La Cruz - Enlace Orotava - Aeropuerto Norte
- Línea 40 Costa Adeje - Los Cristianos - Aeropuerto Sur

*Líneas Metropolitanas*: 19 líneas circulan en el interior del Área metropolitana de Tenerife uniendo Santa Cruz de Tenerife, San Cristóbal de La Laguna, El Rosario y Tegueste. En esta zona se encuentran los dos principales nodos de transporte de la isla, Intercambiador de Transportes de Santa Cruz de Tenerife y El Intercambiador de Padre Anchieta (La Laguna). *Líneas Urbanas*: TITSA opera el servicio urbano de los municipios de Santa Cruz de Tenerife, San Cristóbal de La Laguna, Puerto de la Cruz, La Orotava, Guía de Isora y Los Realejos. *Línea Tuwawa*: En el año 2021 comenzó la prueba piloto de un servicio de transporte a la demanda en los municipios de Fasnia y Arico. *Otros servicios*: La empresa activa servicios especiales para determinadas celebraciones, actos puntuales o temporadas.

- *Paradas\_Subidas, Paradas\_Bajadas*: Conjunto de paradas utilizadas, habiendo más de una al realizarse transbordo.
- *Líneas\_Viajes*: La línea de la guagua que se toma para ir de una parada a otra.
- *Pasajeros\_Entrada*: Número de pasajeros que suben a bordo en la parada de subida descrita.

### 3.1.1. Uso de software QGIS

Estos datos descritos requieren de un proceso de medición ex-post y la transformación del dato obtenido en variable. Por ejemplo, el origen y el destino del viaje son datos importantes para la obtención de algunas variables como la distancia recorrida en guagua/tranvía, para la medición de los tiempos de viajes, qué líneas de guagua/tranvía se toma, etc. Pero ahora se visualizarán las regiones en las que se trabajará, en donde se localizan las diversas paradas y así tener una noción de como va a fluctuar el transporte de pasajeros entre paradas.

Para ello se utilizará QGIS, un Sistema de Información Geográfica de software libre y de código abierto que permite manejar formatos ráster<sup>2</sup> y vectoriales a través de la biblioteca GDAL (GADL/OGR), así como bases de datos. Algunas de sus características son:

- Soporte para la extensión espacial de PostgreSQL<sup>3</sup>, PostGIS<sup>4</sup>.
- Manejo de archivos vectoriales Shapefile, ArcInfo coverages, MapInfo, GRASS GIS, DXF, etc. En este caso se va a trabajar con archivos shapefile, que se explicarán más adelante.
- Soporte para un importante número de tipos de archivos ráster.

Una de sus mayores ventajas es la posibilidad de usar Quantum GIS como GUI (Interfaz gráfica de usuario) del SIG GRASS, utilizando toda la potencia de análisis de este último en un entorno de trabajo más amigable. QGIS está desarrollado en C++, usando la biblioteca Qt para su Interfaz gráfica de usuario.

Con todo esto, se ha usado QGIS para cartografiar la isla de Tenerife y las respectivas paradas como se ve en la figura (3.1), usando las coordenadas geográficas del conjunto de datos anteriormente descrito.

Además, se ha considerado el resaltar zonas de puertos y aeropuertos, las cuáles nos podran dar información sobre el posible flujo de pasajeros turistas en

<sup>2</sup> Una imagen ráster es una estructura o fichero de datos que representa una rejilla rectangular de píxeles o puntos de color, denominada matriz, que se puede visualizar en un monitor, papel u otro dispositivo de representación.

<sup>3</sup> Sistema de gestión de bases de datos relacional orientado a objetos y de código abierto, publicado bajo la licencia PostgreSQL.

<sup>4</sup> PostGIS convierte al sistema de administración de bases de datos PostgreSQL en una base de datos espacial mediante la adición de tres características: tipos de datos espaciales, índices espaciales y funciones que operan sobre ellos.

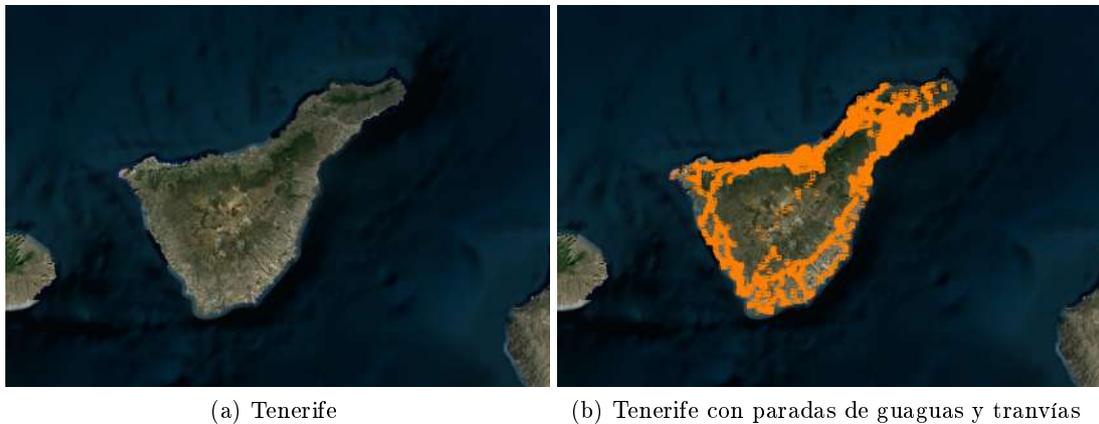


Figura 3.1: Imágen satelital de la isla de Tenerife.

Tenerife, como se ve en la figura (3.2). Con todo esto, al poseer 4.409 paradas distintas, representar gráficamente el flujo de pasajeros entre estas, sería algo engorroso y poco claro de entender. Por ello, mediante el uso de RStudio, se ha reducido el número de paradas, comenzando con dividir la isla de Tenerife en sus respectivos municipios, que están compuestos por un área de terreno delimitado (véase la figura (3.3a)).

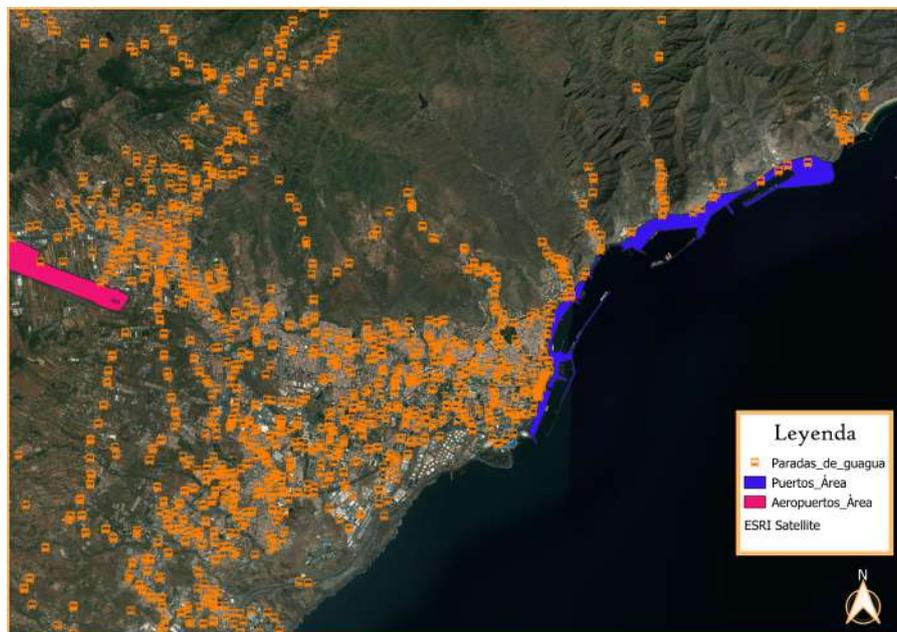


Figura 3.2: Paradas existentes entre los municipios de Santa Cruz de Tenerife y San Cristobal de La Laguna.

### 3.1.2. Archivos SHP

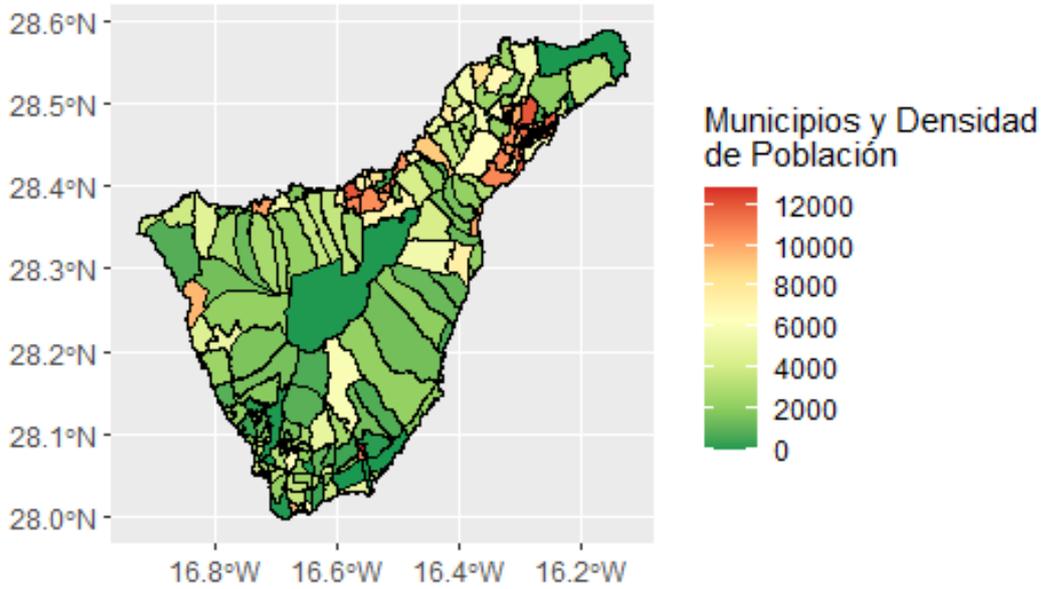
La elaboración descrita anteriormente se ha llevado a cabo leyendo y cargando un conjunto de datos SHAPE proporcionados por TITSA.

Estos archivos denominados *shapefiles* fueron desarrollados por ESRI, una de las primeras y ahora sin duda la mayor empresa comercial de **SIG** (datos de sistema de información geográfica) del mundo. A pesar de ser desarrollados por una empresa comercial, son un formato abierto y los shapefiles pueden ser utilizados (leídos y escritos) por una gran cantidad de aplicaciones de software SIG. Un shapefile es en realidad una colección de archivos, al menos tres de los cuales son necesarios para que el software SIG muestre el shapefile:

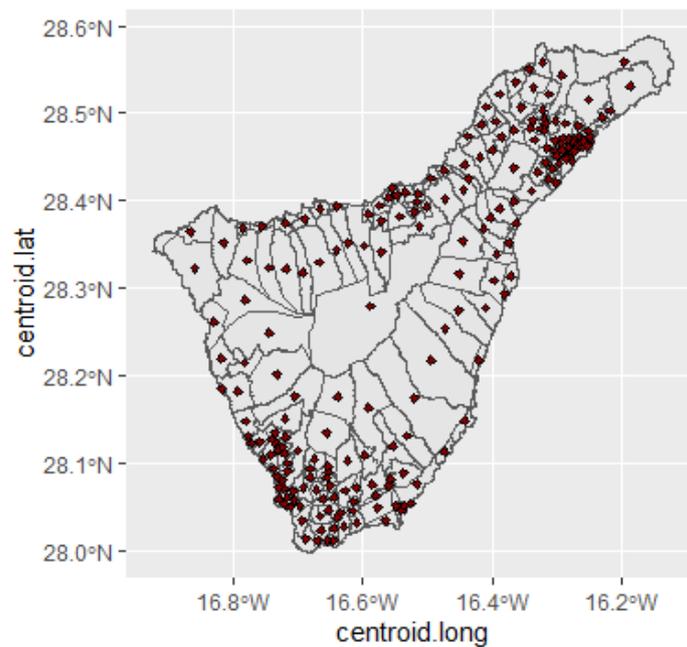
- Extensión de archivo *.shp*: La parte principal del archivo que contiene la geometría característica.
- Extensión de archivo *.shx*: Un archivo de índice que almacena la posición de los ID de las características en el archivo *.shp*
- Extensión de archivo *.dbf*: La parte del archivo que almacena toda la información de atributos asociada a las coordenadas - puede ser el nombre de alguna forma geoétrica o alguna otra información asociada a las características - nótese cómo en un shapefile los datos de atributos se mantienen en un archivo separado.
- Extensión de archivo *.prj*: La parte del archivo que contiene toda la información del sistema de coordenadas. Los datos pueden mostrarse sin una proyección, pero el archivo *.prj* permite que el software muestre los datos correctamente cuando se utilizan datos con diferentes proyecciones. Este archivo no es necesario para mostrar el shapefile.

Además, las entidades geográficas de un shapefile se pueden representar por medio de puntos, líneas o polígonos (áreas). El espacio de trabajo que contiene shapefiles también puede incluir tablas del conjunto de datos, que pueden almacenar atributos adicionales que se pueden vincular a las entidades de un shapefile.

A continuación, se ha considerado los tranbordos existentes entre dos paradas, como líneas de viaje independientes en nuevas filas, que conectan el trayecto entre la parada origen y parada destino iniciales. Posteriormente se calcula el centroide de dichas limitaciones territoriales, como se puede ver en la figura (3.3b), utilizando la función descrita en el Anexo.



(a) Delimitación municipal de la isla de Tenerife con sus respectivas densidades de población mayores de 19 años.

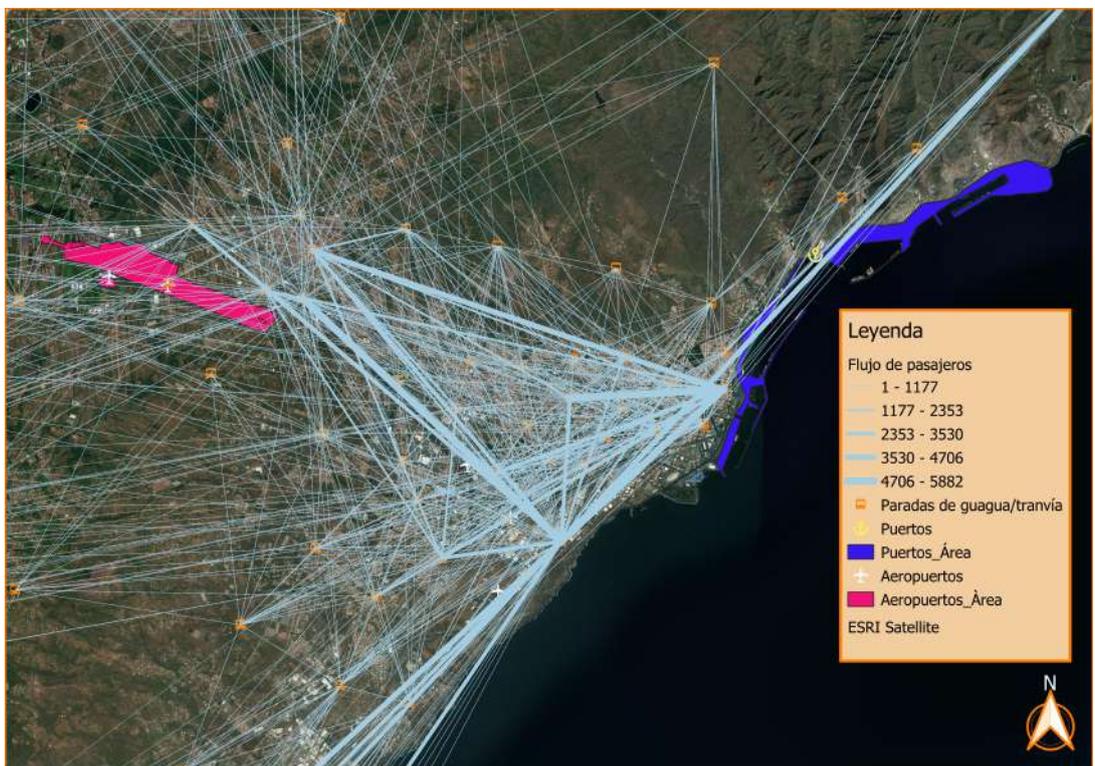


(b) Centroides a los que se asocian las paradas de cada región (Municipio).

Figura 3.3: Visualización y clasificación por distritos y centroides de las paradas de la isla de Tenerife.



(a) Flujo de pasajeros en Tenerife.



(b) Corredor de pasajeros entre Intercambiador de Santa Cruz e Intercambiador de la La Laguna.

Figura 3.4: Flujo de pasajeros en la isla de Tenerife

Con la función anterior se consigue generar el centroide de la figura geométrica que se seleccione. Posteriormente, se ha asignado cada parada de guagua/tranvía al Municipio que perteneciera el centroide cuya distancia a la propia parada fuera la menor. Una vez representada las paradas (centroides) de cada municipio, se ha trasladado lo obtenido a QGIS, para visulizarlo de forma más sencilla y dinámica.

Para ello, con las transformaciones en el conjunto de datos antes mencionadas, se ha obtenido una matriz de origen-destino ( $O - D$ ), utilizando la longitud y latitud de las paradas tanto de subida como de bajada para representar con líneas el flujo de pasajeros, las cuales tendrán un grosor según el número de pasajeros a bordo en la línea de guagua tomada (véase la figura (3.4) ).

## 3.2. Exploración del conjunto de datos en RStudio

A continuación se profundizará en las técnicas de exploración de datos que son útiles para obtener una comprensión de variables numéricas en R. Se presentará aquellos comandos útiles de R que permitan observar los datos de manera eficiente y también crear **Estadísticas Descriptivas**. Para ello, visualizaremos la(s) distribución(es) de nuestros datos mediante la creación de gráficos univariantes. Primeramente se cargará el conjunto de datos `data_od04` (el cuál se ha obtenido luego de haber eliminado valores NaN y renombrado variables para una mejor interpretación de las mismas, con las que se trabajará a partir de ahora) generado anteriormente para crear la cartografía del flujo de pasajeros.

```
1 #Load the data created previously
2 Paradas.Data <- read_excel("data_od04.xlsx")
```

Dicho conjunto de datos está conformado, luego de haber hecho la limpieza en el apartado anterior, por 31 variables y 717.943 datos de pasajeros.

```
1 #List the column headings
2 names(Paradas.Data)
```

### 3.2.1. Introducción al análisis descriptivo

Las estadísticas descriptivas son un medio útil para obtener información rápida sobre un conjunto de datos. RStudio permitirá seleccionar una variable de una ventana desplegable. Antes de continuar, indicar que se ha renombrado las variables del conjunto de datos `data_od04` (ya se habían renombrado después de haber leído y cargado el conjunto de datos iniciales `data_od01`) de la

siguiente forma:

```
1 colnames(data_od01) <- c("profile", "time_on", "date_on", "hour_on", "first_
  stop_on", "last_stop_off", "date_off", "hour_off", "ids_stops_on",
2 "ids_stops_off", "num_trans", "ids_lines", "num_
  passengers")
```

En este caso interesa la variable "num\_passengers" (número de pasajeros).

### Estadística Descriptiva

Valor mínimo	1° cuartil	Media	Mediana	Desviación estándar	3° cuartil	Valor máximo
1,000	1.000	1,345	1	0,690	2,000	24,000

Tabla 3.1: Estadística descriptiva del número de pasajeros que suben a bordo en alguna de las múltiples líneas de transporte existentes.

Como es lógico, podemos ver en la tabla (3.1) que el mínimo número de pasajeros que suben a bordo en la guagua/tranvía es de 1 persona, mientras que el mayor número de pasajeros que han subido en una misma línea es de 24 personas. Aún, teniendo una elevada subida de pasajeros a bordo en muchos casos, la *media* de entrada de personas en guagua/tranvía es de **1.345**. Este bajo valor se debe a que, al tener un gran número de paradas en la isla, existen líneas de transporte que conectan zonas muy poco transitadas, con densidad de población bajas o que no son corredores principales de transporte, ya sea porque no son zonas industriales, turísticas, lugares de trabajo, etc.

Por otro lado, la *mediana* nos da **1**, reflejando que la posición relativa de los datos, es decir, el valor que está en el centro de las observaciones es de una persona por línea de parada tomada. Por tanto, refleja que la frecuencia de que el número de pasajero sea 1 en la subida de paradas es muy alto.

Visualizando la tabla (3.2), se observa que hay **524.882** casos de un solo pasajero en la entrada de una línea de viaje, siendo el **73,07%** de todos los casos, algo muy significativo y que puede refrendar lo explicado anteriormente, incluso llegar a pensar que hay un número excesivo de líneas. La otra reflexión es la de que existen varios corredores de pasajeros de gran afluencia, en las que existen paradas de paso, teniendo cada uno de las personas que suben a bordo la misma parada de destino. La pregunta en hacerse sería la de, ¿son necesarias la existencia de tantas paradas en dichos corredores?

Será más interesante centrarse en municipios con mayores densidades de población (véase la figura (3.5)), más concretamente, en los municipios de Santa

Tabla de frecuencias

Nivel	Frecuencia	Porcentaje sobre el total	Porcentaje acumulado
[1, 2]	682.398	95,048917 %	95,048917 %
(2,3]	24.638	3,431744 %	98,480661 %
(3,4]	6.873	0,957317 %	99,437978 %
(4,5]	2.231	0,310748 %	99,748727 %
(5,6]	912	0,127029 %	99,875756 %
(6,7]	343	0,047775 %	99,923532 %
(7,8]	306	0,042622 %	99,966153 %
(8,9]	149	0,020754 %	99,986907 %
(9, 24]	94	0,013093 %	100,000000 %

Tabla 3.2: Frecuencia del número de pasajeros que suben a bordo en alguna de las múltiples líneas de transporte existentes.

Cruz de Tenerife, San Cristobal de La Laguna, Puerto de la Cruz y Adeje, aunque solo se analizarán en detalle los dos primeros.

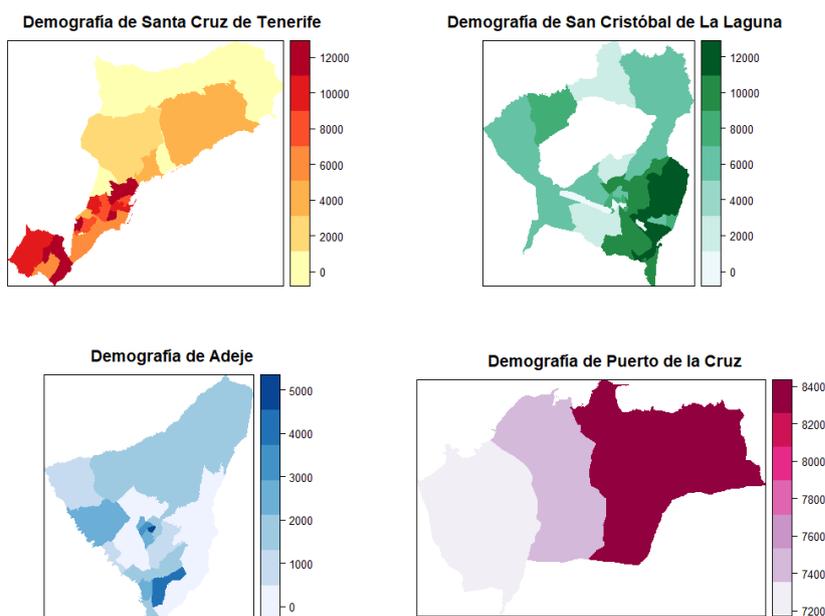


Figura 3.5: Densidad de la población con más de 19 años en los respectivos municipios.

### 3.3. Primer caso de estudio: Santa Cruz de Tenerife

En este caso, se selecciona la parada **“INTERCAMBIADOR DE SANTA CRUZ”**, como centroide de partida y destino, pudiendo visualizar el porcentaje de pasajeros que van a cada una de las partes del municipio, y poder obtener información que aclare que líneas o regiones son más transitadas.

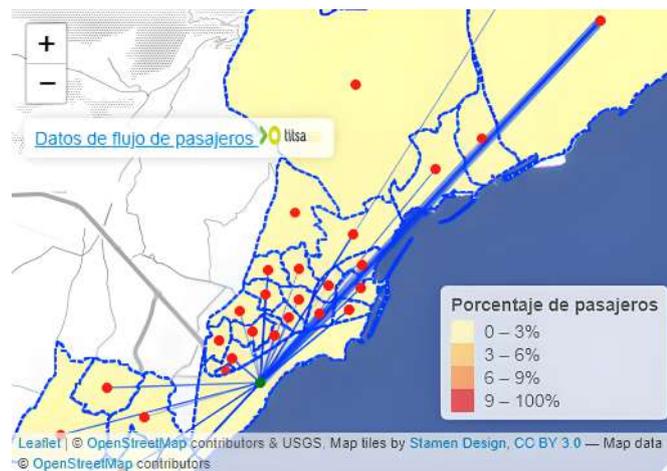


Figura 3.6: Flujo de pasajeros desde el intercambiador de Santa Cruz al resto de distritos del municipio.

Santa Cruz de Tenerife	
<b>Resumen</b>	
Valor mínimo	1,000
1° cuartil	6,000
Media	55.771
Mediana	17,000
3° cuartil	50,000
Valor máximo	2172,000
<b>Conclusiones</b>	
<p>Hay una media de subida de pasajeros alta, ya que se trata del intercambiador ubicado en la capital de la isla, teniendo en cercanías el puerto de Santa Cruz, habiendo mucho movimiento de turistas.</p>	
	

El mayor porcentaje de viajeros se da entre el intercambiador y la parada “Las Teresitas”. Se podría pensar que esto se debe a que en esa zona se encuentra una de las mejores playas de la isla, pero realmente es por la existencia del polígono industrial, ya que son registros de finales de Diciembre y principios de Enero. Hay que destacar que incluso siendo festivo el 25 de Diciembre y el 1 de Enero, hubo mucho flujo por posible motivo laboral, aunque si visualizamos la tabla (3.3)

```
1 data_perfiles <- data_od04 %>% filter ( stop_on.name=="INTERCAMBIADOR_STA.
  CRUZ" , stop_off.name=="LAS_TERESITAS" )
```

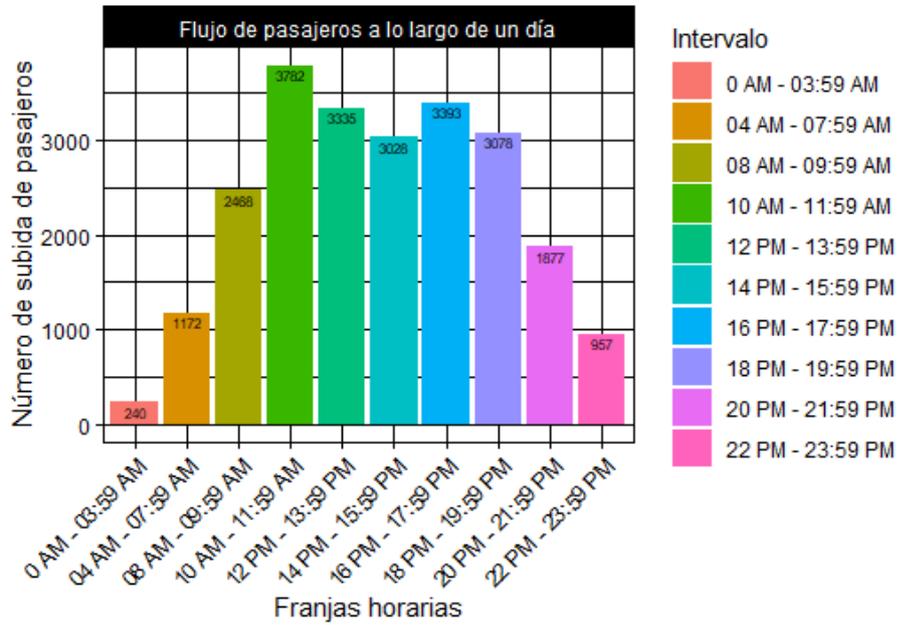
Perfiles		
<i>Desconocido</i> 2	<i>Discapacitado C</i> 26	<i>Discapacitado General</i> 9
<i>Empleado TITSA</i> 1	<i>Estudiante Universitario</i> 1	<i>Familia Numerosa</i> 6
<i>Familiar personal TITSA</i> 9	<i>General</i> 1406	<i>Joven menor de 30</i> 278
<i>Jubilado C</i> 31	<i>Mayor de 65</i> 38	<i>Niño menor de 10</i> 35
<i>Residente Canario</i> 168	-	-

Tabla 3.3: Distintos tipos de bonos utilizados por los pasajeros.

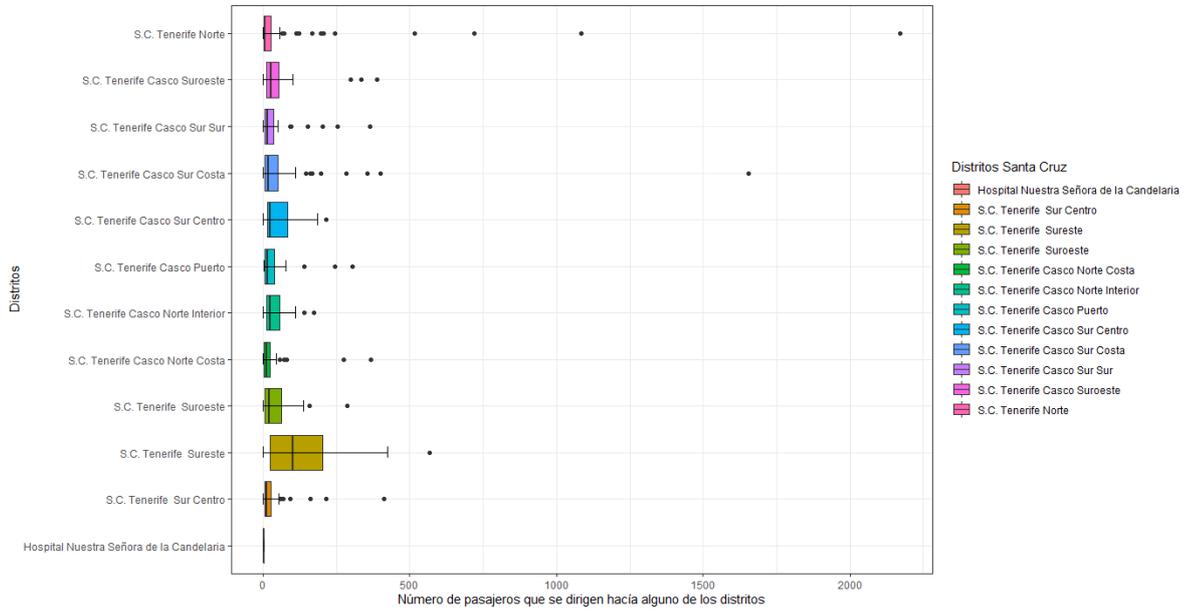
podría ser interesante la existencia de un bono específico para aquellas personas que trabajan en el entorno industrial del municipio, ya que el mayor número de pasajeros viajan sin ningún tipo de bono (General); y así promover el transporte público y evitar las emisiones de gases producidas por el continuo flujo de coche privados.

Tanto en el histograma (figura (3.7a)) como en el del otro municipio a analizar, se visualiza que la distribución de pasajeros que toman el transporte público es más elevada en las horas cercanas a mediodía, ya sea por la finalización de turnos laborales o por estar en época de Navidad y sean momentos de compras. Hay que destacar que en las dos últimas horas del día hubo 957 personas que tomaron alguna de las guaguas que parten desde el intercambiador al resto del municipio. Se intuye que esto estará relacionado por lo anteriormente dicho, de estar en días de Navidad, donde no hay clases escolares o se tienen vacaciones de trabajo, y la gente sale a cenar o ha dar un paseo sin importar llegar tarde a casa.

Podemos ver primeramente en la tabla (3.7b) que los datos que nos arroja el *boxplot* están sesgados, al no ser simétrica la división de las cajas por las



(a) Momentos del día donde se hace uso del transporte publico.



(b) Distritos destinos de los pasajeros que parten del intercambiador de Santa Cruz.

Figura 3.7: Gráficos univariantes

lineas verticales. También se atisba que el porcentaje de subida de pasajeros por debajo del 25% es muy similar entre los diferentes distritos. Sin embargo, el porcentaje de observaciones de subida de pasajeros por debajo del 75% es

mucho más elevado en el destino *Sureste* de Santa Cruz, es decir, hay mucho más flujo de personas hacia esa zona, con un solo valor atípico, pudiendo reflejar que los datos están distribuidos de igual forma (el número de pasajeros a bordo hacia esa zona suele ser idéntico).

A continuación se visualizará en la figura (3.8) como varía el flujo de pasajeros, pero en este caso, hacia otros municipios, seleccionando aquellos corredores donde han subido a bordo de una parada a otra más de 1.000 personas.

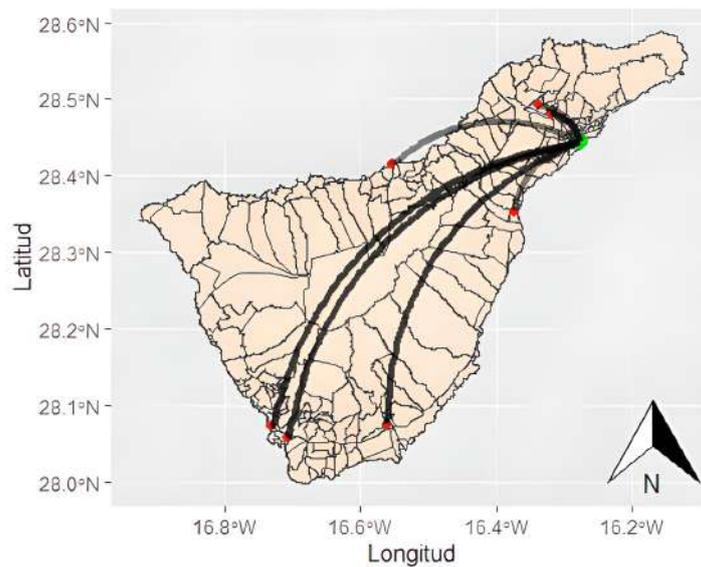


Figura 3.8: Flujo de pasajeros desde el intercambiador de Santa Cruz al resto de la isla.

### 3.4. Segundo caso de estudio: San Cristobal de La Laguna

En este caso, se selecciona la parada *“INTERCAMBIADOR LAGUNA”*, como centroide de partida y destino, pudiendo visualizar el porcentaje de pasajeros que van a cada una de las partes del municipio, y poder obtener información que aclare que líneas o regiones son más transitadas.

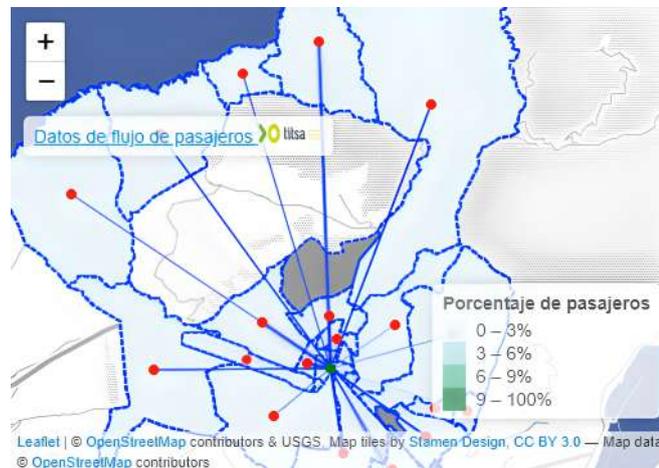
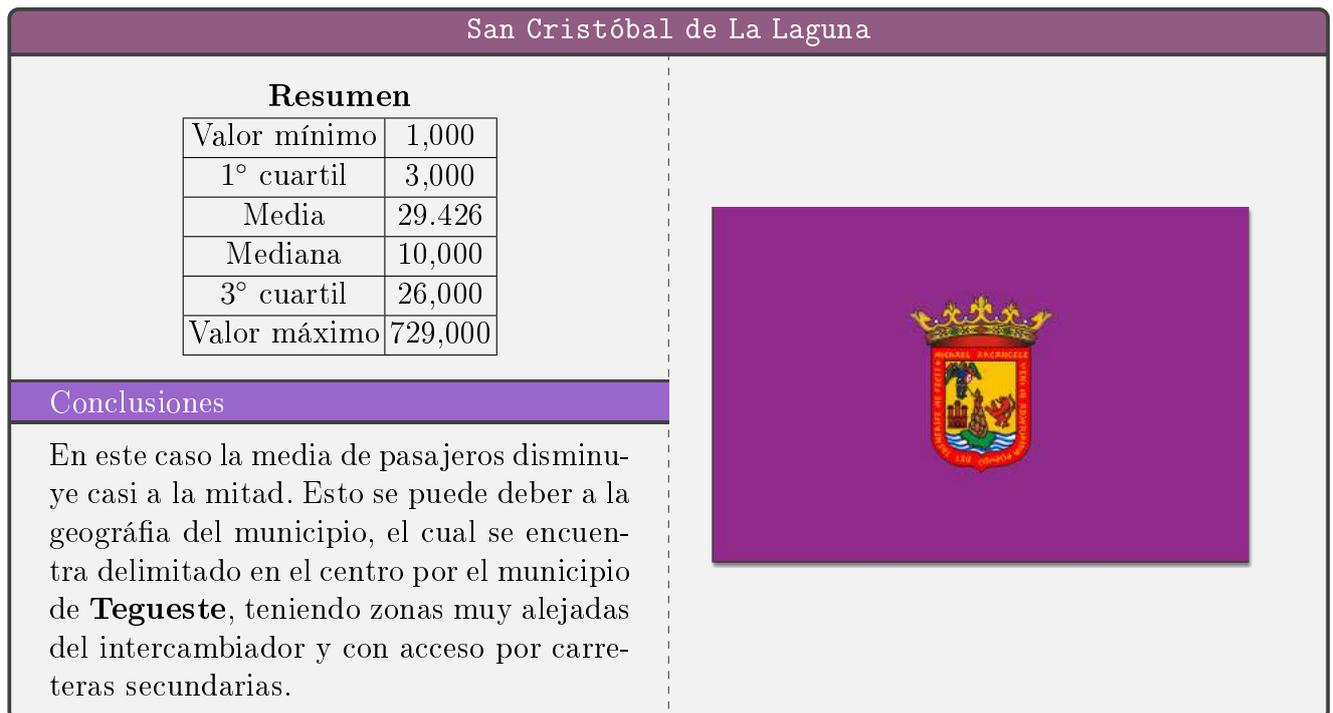
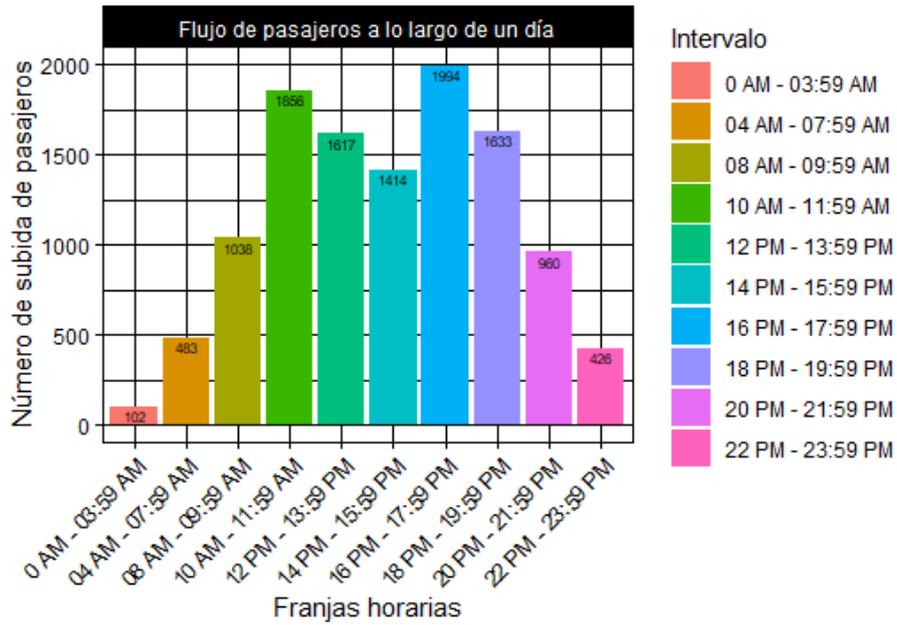


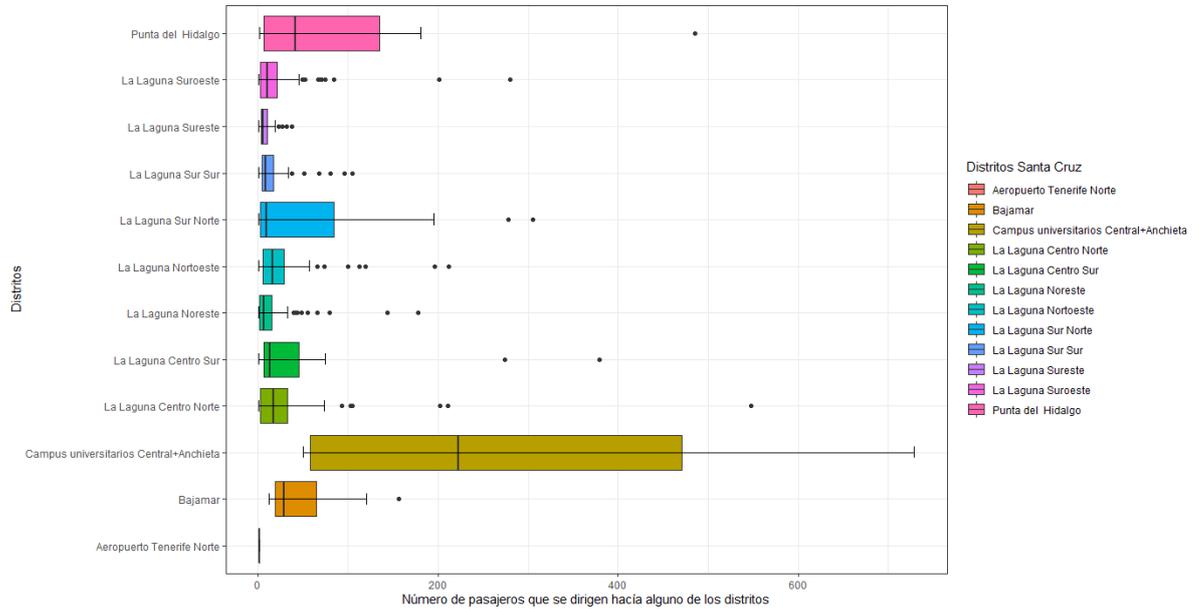
Figura 3.9: Flujo de pasajeros desde el intercambiador de San Cristobal de La Laguna al resto de distritos del municipio.



A destacar, que en el flujo de pasajeros hacia el **aeropuerto**, aún produciéndose en época de festividad (Semana de Navidad), solo representa el 4,501 % del total de trayectos desde el intercambiador. Se puede entrever que la mayoría de pasajero son turistas, al tenerse 278 personas que han viajado sin bono, de un total de 440.



(a) Momentos del día donde se hace uso del transporte publico.



(b) Distritos destinos de los pasajeros que parten del intercambiador de Santa Cruz.

Figura 3.10: Gráficos univariantes

Sería interesante la existencia de un bono compuesto por 15 viajes, ya que podría facilitar el organigrama de viaje de un turista, debido a que los dos bonos turísticos existentes en TITSA a día de hoy, tienen un limite de uso por, en

primer lugar, 24 horas, que no permite el poder transportarte, por ejemplo, a lo largo de un fin de semana, quedándose corto; y por otro lado, un uso de 7 días, que en muchos casos, puede ser algo excesivo en cuanto precio (50 €) y al que puede no llegar a dársele toda la utilidad.

En este caso (véase la figura (3.10a)), el flujo de pasajeros se distribuye de forma más elevada entre las 10:00 y las 12, y en las horas de tarde, disminuyendo el número de subida de pasajeros al mediodía.

A continuación se visualizará en la figura (3.11) como varía el flujo de pasajeros, pero en este caso, hacia otros municipios, seleccionando aquellos corredores donde han subido a bordo de una parada a otra mas de 1.000 personas.

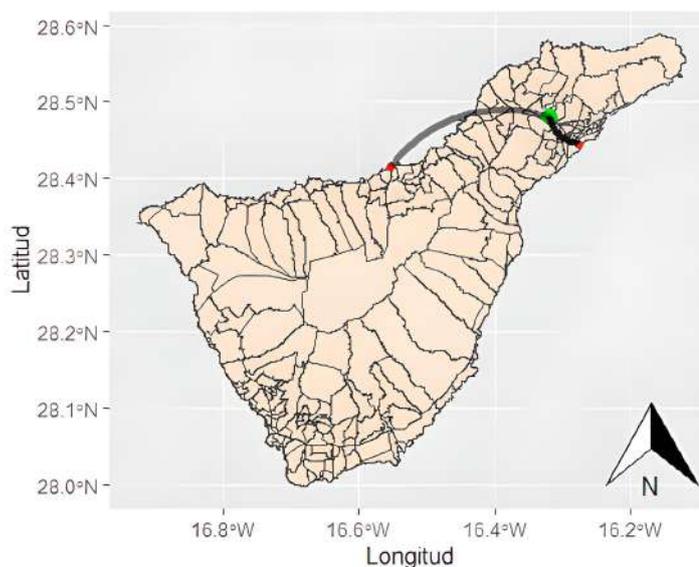


Figura 3.11: Flujo de pasajeros desde el intercambiador de San Cristóbal de La Laguna al resto de la isla.

### 3.5. Estimación de pasajeros mediante modelo de gravedad

Una vez hecho un análisis estadístico descriptivo y espacial de los datos, se utilizará el modelo de gravedad para estimar el número de pasajeros que habrá en las paradas del *Intercambiador de Santa Cruz de Tenerife, San Cristóbal de La Laguna, Adeje y Puerto de la Cruz*. En primer lugar se tiene en la tabla (3.4)

los viajes producidos desde los cuatro intercambiadores, y los viajes atraídos o que llegan a los intercambiadores desde diferentes paradas.

Zona	Producción	Atracción
Intercambiador de Santa Cruz	49.066	71.171
Intercambiador de La Laguna	20.087	31.342
Estación Costa Adeje (T)	14.279	21.495
Puerto de la Cruz	25.286	24.742
Total	108.718	148.750

Tabla 3.4: Producción y Atracción de viajes.

Por otro lado se considera la matriz de costes medios (en minutos) de los viajes entre diferentes paradas o zonas (véase la tabla (3.5)). Por simplicidad, el intercambiador de Santa Cruz se define como la zona 1, el intercambiador de San Cristóbal de La Laguna como la zona 2, el intercambiador de Costa Adeje como la zona 3 y el intercambiador de Puerto de la Cruz como la zona 4.

Matriz de Costes (minutos).				
Zonas	1	2	3	4
1	0	24,36	79,69	49,58
2	23,74	0	108,72	34,91
3	78,70	101,74	0	121,37
4	54,07	45,88	138,52	0

Tabla 3.5: Costes medios (en minutos) de los viajes entre los diferentes intercambiadores.

El primer paso será el ajustar los valores de atracción de las cuatro zonas, debido que al tener un sistema cerrado, la suma de las producciones y atracciones de cada zona deben ser iguales, para poder así calibrar el modelo (exigencia del mismo). Para esto, se ajustan los valores de atracciones para sumar el mismo total que lo producido por las cuatro zonas, mediante el factor:

$$\frac{108.718}{148.750} = 0,7309$$

Ahora multiplicaremos cada valor de atracción por el factor resultante y se obtiene así la nueva tabla **P-A** (véase la tabla (3.6))

A continuación se construye una matriz de los valores  $\exp(-\beta c_{ij})$ , como en la tabla (3.7).

Zona	Producción	Atracción
Intercambiador de Santa Cruz	49.066	52.018
Intercambiador de La Laguna	20.087	22.907
Estación Costa Adeje (T)	14.279	15.710
Puerto de la Cruz	25.286	18.083
Total	108.718	108.718

Tabla 3.6: Producción y Atracción de viajes.

Zonas	$exp(-\beta c_{ij})$			
	1	2	3	4
1	0	0,0875	0,00035	0,0070
2	0,0931	0	0,00001898	0,0305
3	0,00038	0,00003815	0	0,000005358
4	0,0045	0,0102	0,0000009642	0

Tabla 3.7: Valores de la función de disuasión para cada coste de viaje.

Utilizando posteriormente la expresión compacta general del modelo de gravedad:

$$T_{ij} = P_i \frac{A_j \exp(-\beta c_{ij})}{\sum_j A_j \exp(-\beta c_{ij})}$$

y una vez organizado las tablas anteriores, comenzamos a calcular el número de viajes entre las 4 zonas:

$$T_{11} = 0$$

$$T_{12} = P_1 \cdot \frac{A_2 \cdot F_{12}}{A_1 F_{11} + A_2 F_{12} + A_3 F_{13} + A_4 F_{14}} = 46.032, 2633$$

y así con el resto de  $T_{ij}$  de la matriz, hasta obtener la tabla (3.8).

Se comprueba que los parámetros de producción estimadas son casi idénticas a las esperadas, sin embargo, esto no se cumple para los atraídas. Por tanto, se debe realizar la factorización de columnas y filas (en este caso de columnas) hasta que se alcance la solución “convergente”. Se comienza por calcular la tasa de error, que luego se convierte en el factor de columna, que se aplica a todos los viajes en esa columna, lo que garantiza que la conservación de los viajes atraídos se cumpla.

$$FC_i = \frac{A_i}{E_i} \tag{3.1}$$

Viajes estimado entre zonas y totales						
Zonas	1	2	3	4	Total	Objetivo
1	0	46.032,2633	126,2798	2.907,0873	49.065,6304	49.066
2	18.032,2811	0	1,11	2.053,6085	20.086,9996	20.087
3	13.610,5773	601,7305	0	66,7133	14.279,0211	14.279
4	12.654,2037	12.630,98	0,8189	0	25.286,0026	25.286
Total	44.297,0621	59.264,9738	128,2087	5.027,4091	108.717,6537	
Objetivos	52.018	22.907	15.710	18.083		108.718

Tabla 3.8: La matriz del modelo de gravedad resultante con la distribución de la longitud del viaje.

donde  $E_i$  es el valor de atracción estimado de la zona  $i$ . Tenemos así la tabla (3.9)

Factor de columna				
Tráfico de Zonas	1	2	3	4
Atracciones Estimadas	44.297,0621	59.264,9738	128,2087	5.027,4091
Atracciones Objetivos	52.018	22.907	15.710	18.083
FC	1,1743	0,3865	122,5345	3,5969

Tabla 3.9: Tabla con el correspondiente factor de cada columna.

Una vez obtenida la tabla de factores de columnas, realizamos la primera iteración (ver (3.10)).

Viajes estimado entre zonas y totales						
Zonas	1	2	3	4	Total	Objetivo
1	0	17.791,4698	15.473,6322	10.456,5023	43.721,6043	49.066
2	21.175,3077	0	136,0133	7.386,6244	28.697,9454	20.087
3	15.982,9009	232,5688	0	239,9611	16.455,4308	14.279
4	14.859,8314	4.881,8738	100,3435	0	19.842,0487	25.286
Total	52.018,04	22.905,9124	15.709,989	18.083,0877	108.717,0291	
Objetivos	52.018	22.907	15.710	18.083		108.718

Tabla 3.10: Matriz con el número de viajes entre zonas (Primera Iteración).

Con esto se ha conseguido que el valor total de los viajes atraídos sea casi idéntico a las atracciones objetivos. Sin embargo, se ha destruido la conservación de las producciones. Para solucionarlo, se realiza un factor similar al realizado anteriormente, pero por filas (véase la tabla (3.11)).

Una vez obtenida la tabla de factores de filas, realizamos la segunda iteración (véase (3.12)).

Factor de columna				
Tráfico de Zonas	1	2	3	4
Producciones Estimadas	43.721,6043	28.697,9454	16.455,4308	19.842,04871
Producciones Objetivos	49.066	20.087	14.279	25.286
FF	1,1222	0,6999	0,8677	1,2744

Tabla 3.11: Tabla con el correspondiente factor de cada columna.

Viajes estimado entre zonas y totales						
Zonas	1	2	3	4	Total	Objetivo
1	0	19.965,5874	17.364,51	11.734,2869	49.064,3843	49.066
2	14.820,5979	0	95,1957	5.169,8984	20.085,692	20.087
3	13.868,3631	201,7999	0	208,2142	14.278,3772	14.279
4	18.937,3691	6.221,46	127,8778	0	25.286,7069	25.286
Total	47.626,3301	26.388,8473	17.587,5835	17.112,3995	108.715,1604	
Objetivos	52.018	22.907	15.710	18.083		108.718

Tabla 3.12: Matriz con el número de viajes entre zonas (Segunda Iteración).

Se realiza este proceso iterativo hasta obtener unos totales de producción y atracción que estén suficientemente cercanos a los valores objetivos de cada zona. Es decir, se necesita garantizar que los valores calculados estén dentro de un determinado umbral del valor objetivo (1-2%). Teniendo en cuenta esto, y luego de 24 iteraciones, se obtiene la tabla (3.13).

Viajes estimado entre zonas y totales							
Zonas	1	2	3	4	Total	Objetivo	Diferencias %
1	0	19.258,3647	15.615,7993	14.191,6632	49.065,8272	49.066	0.0004 %
2	16.262,0083	0	51,889	3.773,5148	20.087,4121	20.087	0.0021 %
3	14.030,3599	108,3258	0	140,1071	14.278,7858	14.279	0.0015 %
4	21.649,7282	3.574,8562	72,6234	0	25.297,2078	25.286	0.0443 %
Total	51.942,0964	22.941,5467	15.740,3117	18.105,2851	108.729,2629		
Objetivos	52.018	22.907	15.710	18.083		108.718	
Diferencias %	0.1459 %	0.1508 %	0.1929 %	0.1288 %			

Tabla 3.13: Matriz con el número de viajes entre zonas (Vigesimo cuarta Iteración).

Se consigue así la estimación de los viajes entre los cuatro intercambiadores en diferentes rangos de tiempo (véase la tabla (3.14)).

Rangos (minutos)						
Costes	[1, 25]	[25, 50]	[50, 75]	[75, 100]	[100, 125]	[125, 150]
Viajes	35.520,0373	21.540,0342	21.649,7282	29.646,1592	300,3219	72,6234

Tabla 3.14: Número de viajes según el tiempo del trayecto entre intercambiadores.

---

## Conclusiones

En este trabajo se ha abordado teóricamente el qué es la demanda de transporte y como se caracteriza, profundizando en el modelo de cuatro etapas, el más conocido y utilizado en la literatura hasta la fecha. Para ello se han definido las características de los problemas de transporte, así como las cuestiones y fundamentos necesarios para enfocar en una dirección u otra el proceso de modelaje.

Entre los conceptos más básicos pero a la vez importantes, se encuentran la demanda y oferta de transporte, que ayudan a entender como puede variar el flujo de pasajeros entre, ya sean zonas (de forma agregada), o paradas de diversos transportes público (de manera desagregada).

Dentro del proceso de la distribución de viajes, se ha decidido seleccionar el *modelo de gravedad*, por su sencillez, y viabilidad para ajustarse mejor al conjunto de datos que ha sido proporcionado por la entidad TITSA, la cuál contaba con una matriz de orígenes y destinos de cada una de las paradas de la isla de Tenerife; así como el tiempo de todos y cada uno de los trayectos, que se recogían en el segundo conjunto de datos. Además, parte con la ventaja de poder calibrarse y así fijar sus parámetros para que el modelo de viaje del tiempo (día, mes, año) base esté bien representado.

Ha sido eficaz apoyarse en la herramienta QGIS para visualizar el flujo de pasajeros entre paradas, siendo un software a tener en cuenta en futuras investigaciones en la línea de este trabajo, aportando una dinámica y eficacia a la hora de cartografiar, como en este caso, no tan solo diversos municipios, sino zonas de importancia como aeropuertos o puertos.

Se ha utilizado el lenguaje de programación estadístico R, para manipular los datos obtenidos, realizar la estadística descriptiva y llevar a cabo un análisis espacial del conjunto de datos, permitiendo con esto ver que podría afectar al flujo de pasajeros, así como la horas de mayor transito, siendo estas las horas

de media mañana y tarde; y desde qué distrito de los diferentes municipios de la isla producían mayor número de pasajeros de subida en un transporte.

Los datos arrojados por el modelo anteriormente descrito fueron satisfactorios, remarcando el que el número de pasajeros en los viajes que duraban como máximo 25 minutos era mucho más elevado frente al resto de franjas de tiempo de duración de los trayectos. Sería interesante en futuras investigaciones poseer un conjunto de datos con variables socioeconómicas que permitan visualizar el motivo de los viajes de cada uno de los pasajeros (mediante encuestas), y correlacionarlo tanto con la duración de los viajes, como el destino del viaje, así como incluso la edad de cada uno de los pasajeros.

# A

---

## Apéndice

### A.1. Código en lenguaje R para el análisis del conjunto de datos proporcionados por TITSA

Lectura y carga del archivo shapefile:

```
1 ##### load zone maps
2 tene_zones <- read_sf('Zonificacion2.shp',)
3 Encoding(tene_zones$Nombre) <- "latin1"
4 Encoding(tene_zones$NombreEDM) <- "latin1"
5 Encoding(tene_zones$Nmun) <- "latin1"
6 tene_zones$Nombre <- enc2utf8(tene_zones$Nombre)
7 tene_zones$NombreEDM <- enc2utf8(tene_zones$NombreEDM)
8 tene_zones$Nmun <- enc2utf8(tene_zones$Nmun)
```

Clasificación del número de transbordos realizados por un pasajero, en viajes independientes, así como la limpieza y posterior ordenación para facilitar el uso de los datos:

```
1 ##### load OD data
2 data_od01 <- read_excel("datos_od.xlsx")
3 colnames(data_od01) <- c("profile", "time_on", "date_on", "hour_on", "first_
4   stop_on", "last_stop_off", "date_off", "hour_off", "ids_stops_on",
5   "ids_stops_off", "num_trans", "ids_lines", "num_
6   passengers")
7 data_od02 <- data_od01 %>% filter(num_trans >= 0) %>%
8   separate(ids_lines, c("line-01", "line-02", "line-03", "line-04"),
9   sep="," , remove=FALSE) %>%
10  separate(ids_stops_on, c("stop_on-01", "stop_on-02", "stop_on-03", "stop_
11  on-04"),
12  sep="," , remove=FALSE) %>%
13  separate(ids_stops_off, c("stop_off-01", "stop_off-02", "stop_off-03", "
14  stop_off-04"),
15  sep="," , remove=FALSE)
16 data_od03 <- data_od02 %>%
17 pivot_longer(cols = c(starts_with("stop_on-"), starts_with("stop_off-")),
```

```

16           starts_with("line-")),
17           names_to = c("type", "trip_num"), names_sep = "-", values_
           to = "values") %>%
18 pivot_wider(names_from = "type", values_from = "values" ) %>%
19 filter(!is.na(stop_on)) %>%
20 mutate(stop_on=as.numeric(trimws(stop_on)), stop_off=as.numeric(trimws(
           stop_off)), line=as.numeric(trimws(line))) %>%
21 mutate(year=lubridate::year(date_on), month=lubridate::month(date_on),
           label=TRUE, abbr=FALSE), day=lubridate::day(date_on))
22
23 data_od04<-data_od03 %>% left_join(data_stops.sf2 %>%
24                                 as.data.frame() %>%
25                                 select(zone, name.zone, stop, name.stop
26
27                                 ,
28                                 centroid.long, centroid.lat ,
29                                 muni) ,
30                                 by=c("stop_on"="stop")) %>%
31 rename(stop_on.zone=zone ,
32         stop_on.namezone=name.zone ,
33         stop_on.name=name.stop ,
34         stop_on.long=centroid.long ,
35         stop_on.lat=centroid.lat ,
36         stop_on.muni=muni) %>% left_join(data_stops.sf2 %>%
37                                 as.data.frame() %>%
38                                 select(zone, name.zone ,
39
40                                 ,
41                                 centroid.long ,
42                                 centroid.lat , muni) ,
43                                 by=c("stop_off"="stop")) %>%
44 rename(stop_off.zone=zone ,
45         stop_off.namezone=name.zone ,
46         stop_off.name=name.stop ,
47         stop_off.long=centroid.long ,
48         stop_off.lat=centroid.lat ,
49         stop_off.muni=muni)

```

### Cálculo del centroide de cada distrito:

```

1 st_centroid_within_poly <- function (poly) {
2
3   # check if centroid is in polygon
4   centroid <- poly %>% st_centroid()
5   in_poly <- st_within(centroid, poly, sparse = F) [[1]]
6
7   # if it is, return that centroid
8   if (in_poly) return(centroid)
9
10  # if not, calculate a point on the surface and return that
11  centroid_in_poly <- st_point_on_surface(poly)
12  return(centroid_in_poly)
13 }

```

---

## Bibliografía

- [1] Ortuzar, J. de D., & G. Willumsen, L. (2011). *Modelling Transport*. (4.a ed.) John Wiley & Sons Ltd.
- [2] Ben-Akiva, M. E., & Lerman, S. R. (1997). *Discrete Choice Analysis: Theory and Application to Travel Demand*. (7.a ed.) MIT Press.
- [3] Manheim, M.L. (1979). *Fundamentals of Transportation Systems Analysis*. MIT Press.
- [4] Hutchison, C.S. (1974). *Laboratory Handbook of Petrographic Techniques*. John Wiley & Sons Inc.
- [5] Stopher, P., & Meyburg, A.H. (1975). *Urban Transportation, Modeling and Planning*. Rowman & Littlefield
- [6] Wilson, A.G. (1974). *Urban and Regional Models in Geography and Planning*. John Wiley & Sons Inc.
- [7] Daly, A.J., van der Valk, J. & van Zwam, H.P.H. (1983). *Application of disaggregate models for a regional transportation study in the Netherlands*. In P. Baron and H. Nuppenau (eds.), *Research for Transport Policies in a Changing World*. SNV Studiengesellschaft Nahverkehr, Hamburg.
- [8] Wilson, A.G. (1970). *Entropy in Urban and Regional Modelling*. Pion.
- [9] Hensher, D. A. (2013). Decisión de autorías. En André De Palma, Robin Lindsey, Emile Quinet y Roger Vickerman. *Valuation of travel time savings*. En *A Handbook of Transport Economics*. (pp. 135–159). Edward Elgar Publishing Ltd.
- [10] Walker, J. L., & Ben-Akiva, M. E. (2013). Decisión de autorías. En André De Palma, Robin Lindsey, Emile Quinet y Roger Vickerman. *Advances in discrete choice: mixture models*. En *A Handbook of Transport Economics*. (pp. 160–187). Edward Elgar Publishing Ltd.
- [11] Fosgerau, M., & de Palma, A. (2013). Decisión de autorías. En André De Palma, Robin Lindsey, Emile Quinet y Roger Vickerman. *Dynamic traffic modeling*. En *A Handbook of Transport Economics*. (pp. 188–212). Edward Elgar Publishing Ltd.

- [12] Rawoof Pinjari, A., & C. R. Bhat (2013). Decisión de autorías. En André De Palma, Robin Lindsey, Emile Quinet y Roger Vickerman. *Activity-based travel demand analysis*. En *A Handbook of Transport Economics*. (pp. 213–248). Edward Elgar Publishing Ltd.
- [13] Williams, H.C.W.L., & Ortuzar, J. de D (1982b). Decisión de autorías. En Juan de Dios Ortuzar y Howard R. Kirby. *Travel demand and response analysis-some integrating themes*. En *Transportation Research* (pp. 345–362). Foreword
- [14] Daly, A.J. (1982b). Decisión de autorías. En Juan de Dios Ortuzar y Howard R. Kirby. *Applicability of disaggregate models of behaviour: a question of methodology*. En *Transportation Research* (pp. 363–370). Foreword
- [15] Daly, A.J., & Ortuzar, J. de D. (1990). Decisión de autorías. En Juan de Dios Ortuzar y Howard R. Kirby. *Forecasting and data aggregation: theory and practice*. En *Traffic Engineering and Control* (pp. 632–643). Foreword
- [16] Lansley, G., & Cheshir, J. (2016). *An Introduction to Spatial Data Analysis and Visualisation in R | CDRC Data*. *CDRC Data*. [en línea]. [Fecha de consulta: 2022]. Disponible en: <https://data.cdrc.ac.uk/dataset/introduction-spatial-data-analysis-and-visualisation-r>.
- [17] R Core Team. *R: A language and environment for statistical computing*. 2019. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>.
- [18] QGIS Development Team. *QGIS Geographic Information System* 2022. Open Source Geospatial Foundation Project. <https://qgis.org>.

# Transport modelling and passenger mobility analysis

Jonathan Déniz Orosa

Facultad de Ciencias • Sección de Matemáticas

Universidad de La Laguna

alu0101265315@ull.edu.es

## Abstract

IN THE TRADITIONAL PROCESS OF TRANSPORT and road planning (Generation, Distribution, Modal Distribution and Allocation), the distribution phase is used to distribute the trips that are produced between the different areas that make up a region, in this case, the island of Tenerife. This paper describes both the theoretical framework and the application of the Mathematical Gravity Model used for trip distribution, as well as its scope and limitations. This model will be calibrated on the basis of the data obtained by TITSA, with the origins and destinations of the journeys made by 717,944 passengers over 9 days, after spatial and statistical analysis of the data set, using the software QGIS and the R statistical programming language and environment.

## 1. Transport planning and modelling

DEMAND FORECASTING is an essential element in the analysis of transport systems analysis:

- It deals with the behaviour of consumers of transport services and facilities.
- Analyses the response of users to changes in the environment around them.

On the other hand, rapid progress in computer technology has led to the development of procedures for predicting trips by origins and destinations and traffic volumes on network links, taking into account:

- Transport demand
- Transport supply
- Transport planning

Finally, the acceptability of modelling, or of a particular modelling approach, within a decision style is very important.

## 2. The structure of the classical transport model

YEARS of experimentation and development have resulted in a general structure that has come to be known as the classical transport model.

The general form of the model is shown in Figure (2). For the modelling of the trip distribution, the gravity model will be used to estimate the total number of trips emanating from an area and those attracted to each area.

The classic transport model is

$$T_{ij} = A_i O_i B_j D_j \exp(-\beta c_{ij}) \quad (1)$$

where:

$O_i$  is the total number of trips originated in zone  $i$ , and  $D_j$  is the total number of trips destined to zone  $j$ .

$\beta$  is a calibration parameter to adjust the model to each specific area.

$c_{ij}$  es el coste de viaje entre la zona  $i$  y  $j$ .

$A_i$  y  $B_j$  the equilibrium factors which are expressed as:

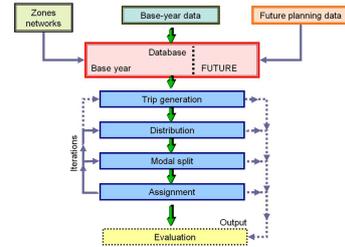


Figure 1: The classic four stage transport model.

$$A_i = \frac{1}{\sum_j B_j D_j \exp(-\beta c_{ij})} \quad \text{y} \quad B_j = \frac{1}{\sum_i A_i O_i \exp(-\beta c_{ij})}$$

## 3. Analysis of passenger traffic in Tenerife

FIRSTLY, the fluctuation of passengers between all existing stops on the island of Tenerife will be visualised and described using Qgis and RStudio.



Figure 2: Passenger corridor between Santa Cruz Interchange and La Laguna Interchange.

With the gravity model, trips between Santa Cruz, San Cristóbal de La Laguna, Adeje and Puerto de la Cruz interchanges are estimated at different time ranges:

	Range (minutes)					
Costs	[1, 25]	[25, 50]	[50, 75]	[75, 100]	[100, 125]	[125, 150]
Trips	35.520,0373	21.540,0342	21.649,7282	29.646,1592	300,3219	72,6234

## References

- [1] Ortuzar, J. de D., & G. Willumsen, L. (2011). *MODELLING TRANSPORT* (4.a ed.) [online]. John Wiley & Sons Ltd.
- [2] Lansley, G., & Cheshire, J. (2016). *An Introduction to Spatial Data Analysis and Visualisation in R — CDRC Data. CDRC Data.* [online]. [Consultation date: 2022]. Available in: <https://data.cdrc.ac.uk/dataset/introduction-spatial-data-analysis-and-visualisation-r>.