Predicting the flow of people for public transportation improvements

Challenge Provider: Porto Digital / Municipality of Porto

Porto Digital was created by the Municipality of Porto in order to promote ICT projects within the context of the city of Porto and its metropolitan area. The Municipality of Porto assigned to Porto Digital the responsibility of maintaining and expanding the telecommunications and IoT infrastructure of the city of Porto (in particular, the optical fibre and Wi-Fi networks, and city sensors), and the task of developing an urban platform to consolidate data in the areas of mobility, environment, energy and civil protection. Accordingly, Porto Digital is a key actor in the design and implementation of smart and sustainable city strategies and policies, which contribute to a better quality of life for the inhabitants, workers, entrepreneurs, visitors and tourists of the city of Porto.

Context

The city of Porto is the 2nd biggest city in Portugal. It spans over an area of 41.42 km^2 and is home to 216 887 inhabitants. In the past years, Porto has been crowned a top destination multiple times and it is known for its vibrant and growing community. However, Porto also tops the charts in the Iberian Peninsula for the worst traffic situation, surpassing Madrid and even Barcelona, and ranking 158th in the world in the 2021 Traffic Index.
There could be many explanations for this:

- The access to the city is being restricted to three bridges and two main highway merge points;
- Difficulty in urban planning due to Porto being an old city built on a hill;
- A public transport system that is unable to meet community demands;

Despite this, Porto is trying to change and to adapt to create a better public transport infrastructure and is working to prioritise soft transportation areas and lanes. Further, Porto is implementing several P+R locations to enable the interchange of private vehicles to public transport.

Further Background Information

Traffic in Porto
https://www.tomtom.com/en_gb/traffic-index/ranking/?country=PT

Public Transport Analysis (PT)
https://www.deco.proteste.pt/familia-consumo/orcamento-familiar/noticias/transportes-publicos-ineficazes-em-6-cidades

Report on the mobility in the Metropolitan Area of Porto - 2017

https://boost.up.pt/news/imob2017/

There is already some work on improving public transport with mobility data that can help to brainstorm for solutions and work as a good basis. For instance:

Previous Challenges of WDL:
https://gitlab.com/worlddataleague/wdl-solutions/-/tree/main/WDL_2021/Stage_1_Public_Transportation

Using Location Data:
Using Data Analytics to Optimize Public Transportation on a College Campus

Using Schedule and Validation Data:
Data-Driven Optimization of Public Transit Schedule

Goal

In this challenge, we want to focus on data to study how to improve the city's public infrastructure to help reduce traffic, improve urban sustainability and quality of life for its citizens.

Sustainable Development Goal

GOAL 11: Sustainable Cities and Communities
Target 11.2.1: Provide access to safe, affordable, accessible and sustainable transport systems for all.

Outcome

Create a model that predicts the in and outflow of people short-term (days) and long-term (months) to and from the municipality of Porto.
By comparing the current and future in/outflow and the entry/exit validation of public transport, identify which routes need current and future interventions and attention from the public office.

Available Resources

All the data resources can be found here:
https://wdl-data.fra1.digitaloceanspaces.com/porto/InterMunicipality.zip

As a reminder, you can also use any data that is open, free and legally available.

The following list of resources is available for you to use:

- Entry and Exit validation data (TIP) from public transportation in the Metropolitan Area of Porto (AMP);

- Origin-Destination (OD) matrices of Movement of People from/to the Porto Metropolitan Area;
- GTFS from Porto's Metro and Public Bus System (STCP);

The description of the datasets can be found in the dictionary.

Tips

- Start by defining what will be the end product and what is it going to answer;
- Don't forget to explore the data in visual and geographical terms;
- Consider first defining a metric to measure the quantity of an aspect you are exploring and then addressing the quality of the aspect in a second step. (e.g. what is the ratio of public transport stops per area, vs the ratio of travels between neighbourhoods);
- Google has helpful documentation on the general GTFS dataset format and structure.

Submissions

Deadline: Wed, 22 April 23h59 AoE (Anywhere on Earth)

Don't forget that you will need to submit the solution report (notebook template with the link below) and executive summary (markdown template below). You also need to submit a **3-minute** video summary of your solution.

Solution report template: https://bit.ly/wdl_2022_jupyter_template
Executive summary template: https://bit.ly/wdl_2022_exec_sum