



**HAL**  
open science

## Opening Telecommunication Data for Inclusive Education

Aránzazu San Juan Llano, Thi-Mai-Trang Nguyen, Daniel Rodriguez,  
Matthieu Chardy

► **To cite this version:**

Aránzazu San Juan Llano, Thi-Mai-Trang Nguyen, Daniel Rodriguez, Matthieu Chardy. Opening Telecommunication Data for Inclusive Education. NetMob 2023, NetMob 2023 Data Challenge, Oct 2023, Madrid, Spain. , 2023. hal-04256613

**HAL Id: hal-04256613**

**<https://hal.science/hal-04256613v1>**

Submitted on 24 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

# OPENING TELECOMMUNICATION DATA FOR INCLUSIVE EDUCATION

ARÁNZAZU SAN JUAN LLANO<sup>1,2</sup>, THI-MAI-TRANG NGUYEN<sup>2,3</sup>, DANIEL RODRÍGUEZ<sup>4</sup> AND MATTHIEU CHARDY<sup>1</sup>

<sup>1</sup>Orange Innovation, France

<sup>2</sup>LIP6 - Sorbonne Université, CNRS, France

<sup>3</sup>L2TI - Université Sorbonne Paris Nord, France

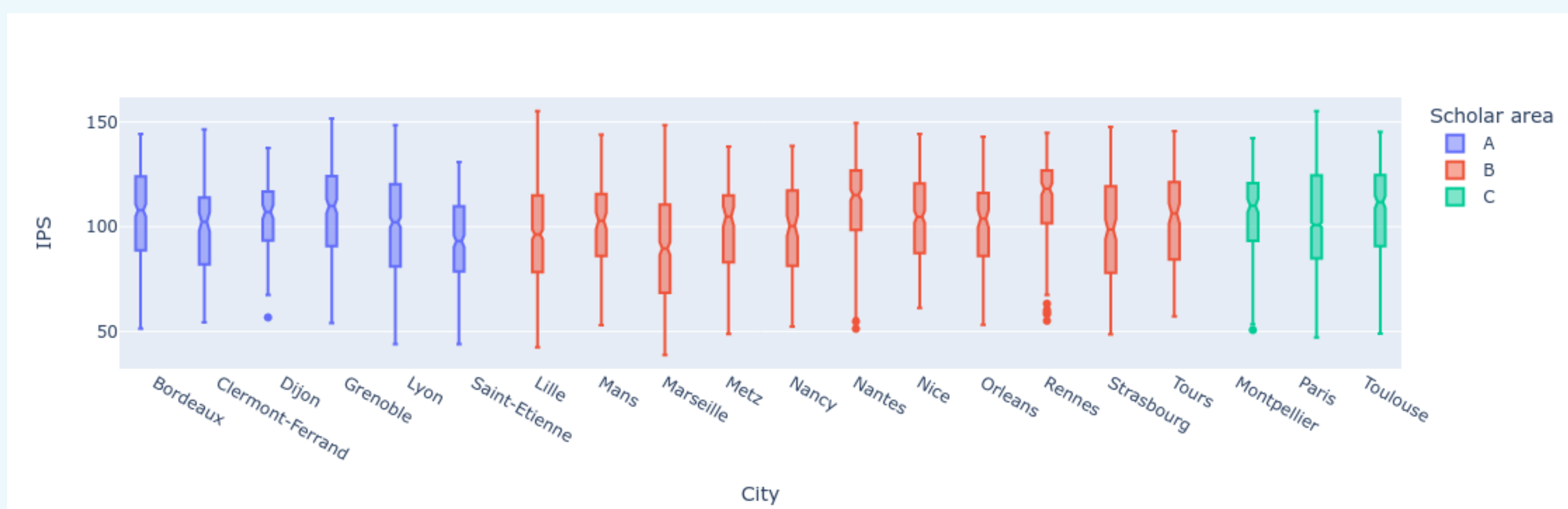
<sup>4</sup>Universidad de Alcalá, Spain

## ABSTRACT

In this poster, we present the main stages of our research in the context of the NetMob 2023 Data Challenge. Following our global research work in cloud continuum infrastructures for inclusive education, we aim to evaluate the impact of telecommunication infrastructures on digital access from an open data perspective in education. We will analyze the French National Education use case crossing the 4G dataset provided by the French telecommunications operator Orange with open datasets for inclusive education in order to propose the first definition of a new Inclusiveness Education KVI (Key Value Indicator) compatible with the Hexa-X 6G KVI definition.

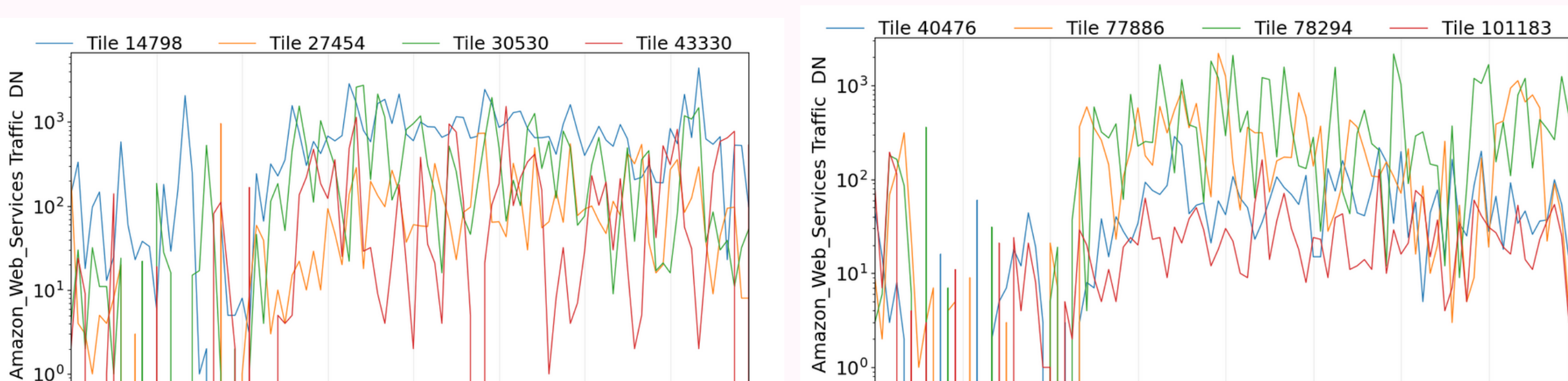
## THE FRENCH NATIONAL EDUCATION USE CASE

- ▶ Define a new Inclusiveness Education KVI extending the Student's Digital Opportunity model proposed by Jim, Evans, and Grant [1] to the mobile network and add other soft indicators emerging from intrinsic peculiarities of the French National Education use case, especially regarding inclusion scope.
- ▶ Work with the open data repository of the French government dedicated to education [2].
- ▶ Cross some open datasets of the education domain with the provided Orange dataset [3].
- ▶ Explore the Social Position Index of schools (named after the French acronym, IPS) of different academic stages (primary, middle, and high school) and the IPS of EREA units from the academic year 2018-2019.



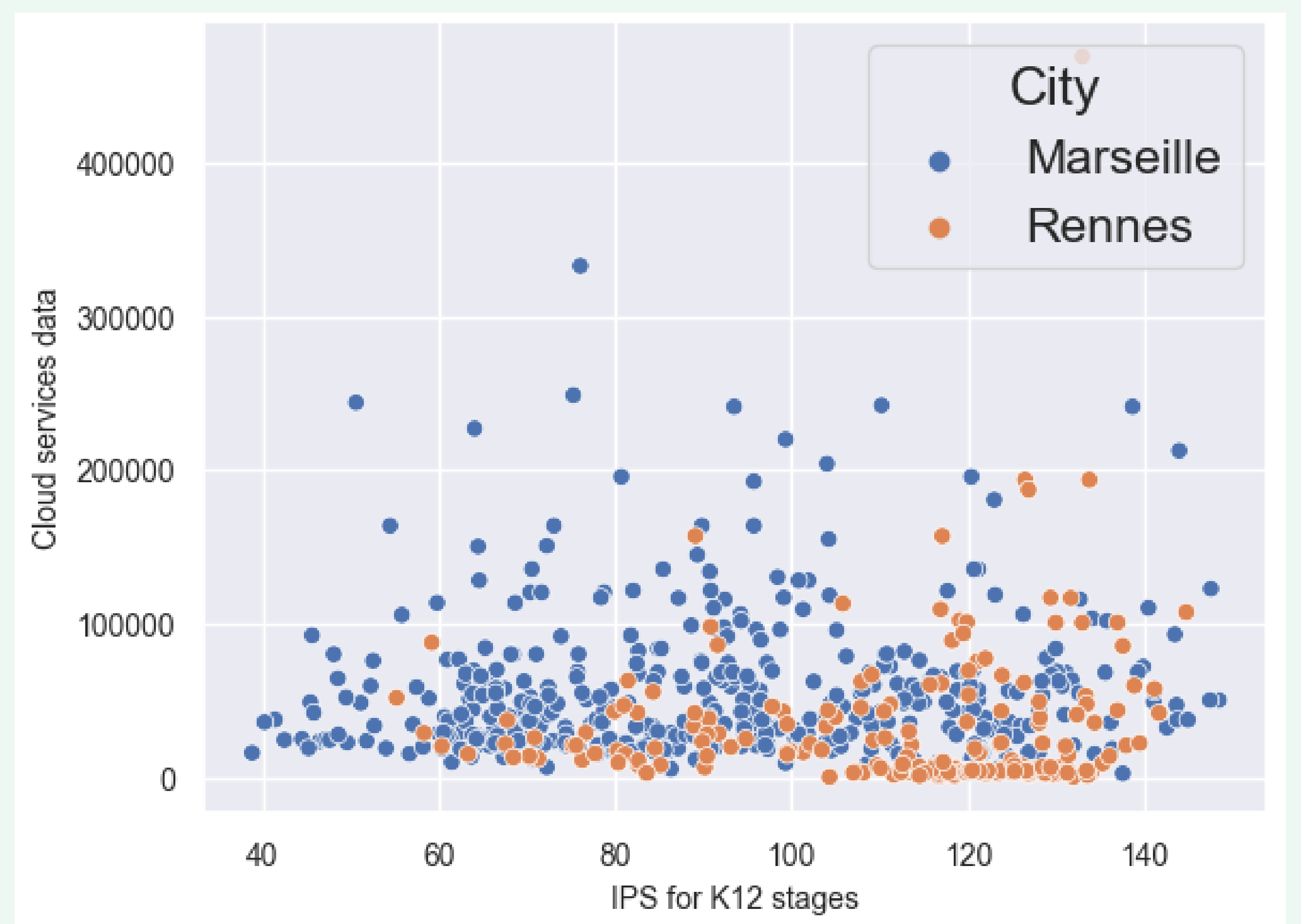
## EXPLORING TRAFFIC ON MIN-MAX IPS TILES

- ▶ Starting the study with the primary schools.
- ▶ The city with the biggest IPS standard deviation is Marseille.
- ▶ The city with the smallest IPS standard deviation is Rennes.
- ▶ Merge the geographical data of schools with the geographical data of every city to give the tile identifier associated with every school.
- ▶ Compute the minimum and maximum IPS school for Marseille and Rennes.
- ▶ With the primary key of schools (UAI), compute the tiles associated with the schools with the min and max IPS of Marseille and Rennes.
- ▶ Compute the traffic for Amazon Web Services for a working day on these min-max tiles.



## EXPLORING TRAFFIC ON ALL SCHOOL TILES

- ▶ Extending the study from the primary schools to all stages (Primary, Middle, and High), EREA schools, and eight cloud services.
- ▶ Merge the traffic file of every city with the aggregated IPS schools file of all stages and combine the traffic of Marseille and Rennes for a working day.
- ▶ Observe the relation between the IPS and the traffic with the correlation matrix and the scatter plot.
- ▶ In the observed period, there is no correlation between the IPS and the data consumption.



## FUTURE WORK

- ▶ Continue to evaluate the relationship between economic factors like the IPS for all applications.
- ▶ Continue to study the temporal and spatial traffic patterns during academic holiday periods to characterize the mobile traffic for future mobile infrastructures for inclusive education.
- ▶ Evaluate the relationship between 4G network availability and data usage for a set of applications equivalent to educational applications of a traditional central cloud infrastructure.
- ▶ Create a first version of our inclusiveness KVI in order to estimate the digital divide across different geographical levels. Improve the accuracy of this inclusiveness KVI, taking into consideration other inclusion open data.
- ▶ Predict the traffic demand for a mobile education infrastructure with Machine learning (ML) based forecasting models.

## REFERENCES

- [1] Cary Jim, Sarah Evans, and Alison Grant. "Multidimensional Approaches to Illustrate to Digital Divide among K-12 Students". In: (July 15, 2021). DOI: 10.29173/ias18282.
- [2] Explore — Éducation Nationale - Accueil. URL: <https://data.education.gouv.fr/explore/?sort=modified>.
- [3] Orlando E. Martínez-Durive, Sachit Mishra, Cezary Ziemlicki, Stefania Rubrichi, Zbigniew Smoreda, and Marco Fiore. *The NetMob23 Dataset: A High-resolution Multi-region Service-level Mobile Data Traffic Cartography*. May 11, 2023. arXiv: 2305.06933. URL: <http://arxiv.org/abs/2305.06933>. preprint.