

# Educational data science: Monitoring learning technologies in primary schools

Natalia da Silva

natalia.dasilva@fcea.edu.uy - natydasilva.com - @pacocuak

IMS-International Conference on Statistics and Data Science (ICSDS)  
Florence, Italy

December-2022



FACULTAD DE  
CIENCIAS ECONÓMICAS  
Y DE ADMINISTRACIÓN

IESTA

INSTITUTO  
DE ESTADÍSTICA



UNIVERSIDAD  
DE LA REPÚBLICA  
URUGUAY

## Data science project, context and data

- Context:** The One Laptop per Child (OLPC) program aims to improve learning in the poorest regions of the world though providing laptops to children for use at school and home. Uruguay implemented this initiative in 2007 and now every children in public education system has a laptop or tablet with specific content to learn.

If we are doing data science, we need DATA:

- DATA:** Students and teachers use of Learning Management Systems (LMS) from primary schools in Uruguay from 2018 to 2021 (for 4th to 6th grades).

Year	N	Students	Techerss
2018	1.912801	98.054	3801
2019	2.652856	103.168	4474
2020	7.403004	109.019	5477
2021	8.607445	119.065	5299

- LMS or learning platforms like Moodle, Blackbord etc, generate a large volume of student and teachers data that can be transformed into a key elements for evaluation and improvement of both face-to-face and distance education.
- To transform these data into relevant information for analysis and decision-making it is required to pre-process the data, define relevant summaries, and statistical visualization.

This project has different objectives:

- 1 Tools for evaluation and monitoring the use of Learning Management Systems in primary schools in Uruguay.
- 2 Identify the key drivers in the use of LMS: understand the sources of variability.
- 3 Predictive models: predict the performance in English tests based on LMS students use. Supervised learning methods were used.

Some included Variables:

- Posted comments in foros
- Posted comments in homework
- Finished homework
- Numbers of time enter into the plataform
- ....

$$IE_i = \frac{\delta+1}{\delta} \left( \frac{e^{\pi_i}}{\delta+e^{\pi_i}} - \frac{1}{\delta+1} \right) \quad \pi_i = \sum_j \log(1+x_{ij})$$

Web application using shiny R package. for monitoring LMS, prototype with data from 2015 and 2017. Explore students engagement in different levels of

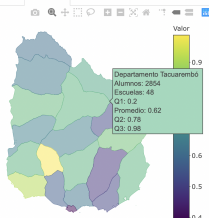
Elija el año  
2017

Elija el índice  
ind\_eng90

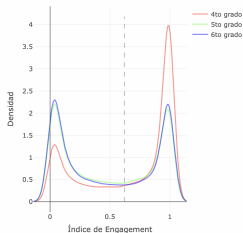
cuantiles - media  
Mediana

Construido con  por   
para  Plan Ceibal financiado por 

Introducción Nacional Escuela Grado Clase Reporte



Mapa de Uruguay para la Mediana de la distribución del índice de engagement en el año 2017



Distribución del índice de engagement en el año 2017 de 4to, 5to y 6to año de primaria

analysis.

<http://164.73.240.157:3838/App-Ceibal/>

Elija el año  
 2017

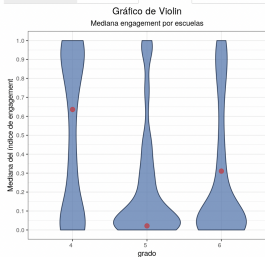
Elija el índice  
 ind\_eng90

Seleccione el departamento  
 Montevideo

Elija la escuela  
 1083

Construido con por para Plan Ceibal  
 financiado por

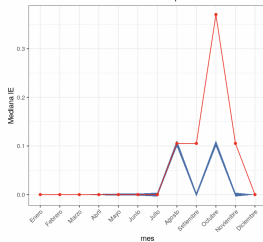
Introducción Nacional Escuela Grado Clase Reporte



El gráfico de violín muestra en el eje vertical la mediana del IE y en el eje horizontal los grados de primaria (4to, 5to y 6to). Representa con color azul la mediana del índice de engagement para cada escuela perteneciente a Montevideo y sus respectivos grados (4to, 5to, 6to). Con color rojo se muestra la mediana del IE de la escuela 1083 para 4to, 5to y 6to.

Grado	Escuela	Departamento	Mediana	Alumnos Totales	Ranking Montevideo	Total Escuelas Montevideo
4	1083	Montevideo	0.64	11	118	259
5	1083	Montevideo	0.02	11	190	262
6	1083	Montevideo	0.31	15	75	245

Gráfico Serie Temporal



El gráfico de serie temporal muestra en el eje vertical la mediana del IE y en el eje horizontal los meses del año 2017. La línea punteada de color rojo muestra la evolución mensual de la mediana del IE de la escuela 1083. Con color azul se muestra la evolución mensual del intervalo de confianza de la mediana de las escuelas pertenecientes a Montevideo.

The use of LMS in the pandemic context increase a lot, more important to improve the computational performance of the app to deal with new an bigger data sets:

- Include new data from 2018 to 2021.
- Define new engagement index to because in 2020-2021 the platform use increase a lot
- Instead of computing everything in R we have a Postgres data base and do the connection with shiny.



Thank you!