# A+ in Math

# If a country wants to progress, its inhabitants must know mathematics.

# "Mathematics is the key and

# Math is the key that allows all kinds of innovations and technological advances

# Kinds

We could say that if our kids know mathematics we can expect a bride future, right? Then

# Achtung!

We have a problem.

# PISA 2018

PISA results show we are falling behind.

# PISA 2018

We are in 16th place! Behind the Netherlands, Sweden and Belgium.

We need to do something.

# What is the key to

We need to figure out the key to improve our results in math scores. Any idea?

# Idea

I have an Idea!

# Most predictive feature of low

If we know the characteristics of students who have low scores in mathematics. Then

# Focus especially on guiding

# we can provide them with greater support during their earlier educational stages.

I already know this features

# ...do you want to know?

Do you want to know them too?

Me too

# Datase/E.D.A./Regression

First, I’m going to explain how did I find them?

# Dataset Kaggle

I took a dataset from Kaggle, unfortunately it is a fictional dataset

# Rows and Columns

It has nineteen thousand rows (after cleaning) and

Fifteen columns

# E.D.A.

The data set contains three different scores of students between 0 and 100 in Math, Reading, Writing.

We can see that in the average Reading and Writing score, **Females are higher** than males but in **math is the opposite.**

# Lunch type

Beside the scores I also have information on Lunch type. There are two types: free/reduce and Standard. Here it **talks the resources** of students family.

# Ethnic group

In the dataset ethnic group are not specified.

# Rest of the…

We also have other categories that we are not going look at in now in detail.

# Regressions

Based in this data I created three models:

One linear regression, one decision tree and one KNN model.

# Linear Regression

# Results

The result of this regression is quiet good, we can see that the R-squared is 0.842.

# Results

The features that have **more impact** on math score are

**Gender Female**, with a negative impact of 11 points of difference in comparison with males.

All the EthnicGroups has **more or less the same results except Ethnic Group E** that seems to have an advantage with positive 5.6 points in comparison to group A

We also see that **children with free lunch have a disadvantage of minus almost five points in** comparison to children with standard lunch (parents that paid for their lunch).

# Decision tree

The Error metric is 0.83-0.82 and that is good. Almost as good as linear regression

# Feature Importance

Here we see that by far the most important feature is the reading score, surprisingly.

This is follow by female gender and free lunch.

# KNN

To complement my analysis I tried a KNN model.

But I saw that the performance was lower than the previous two models.

# Overview

Here are all three regressions and we can see that linear regression is the most **accurate** but there is little difference with decision tree.

# Conclusion

So according with my analysis of this dataset we can conclude that:

We need to ask to the ethinic group E what they are doing to get higher scores in math

And give support to females and students with free lunch type, or in other words, low recourses ramilies.

# Thank you

Please someone call the minister of education.