Subject **Fw: Lab 9+10 Time Series Analysis**
From João Marques <jograma_lk@hotmail.com>
To joaobernardosilva@tecnico.ulisboa.pt <joaobernardosilva@tecnico.ulisboa.pt>, joaofbravo@tecnico.ulisboa.pt <joaofbravo@tecnico.ulisboa.pt>
Date 18-12-2020 15:34

TÉCNICO LISBOA

From: Google Forms <forms-receipts-noreply@google.com>
Sent: Friday, December 18, 2020 3:33 PM
To: joao.g.marques@tecnico.ulisboa.pt <joao.g.marques@tecnico.ulisboa.pt>
Subject: Lab 9+10 Time Series Analysis

Google Forms

Thanks for filling out Lab 9+10 Time Series Analysis

Here's what we got from you:

Lab 9+10 Time Series Analysis

Optional lab evaluation to delivery until December 20th @ 23:59.

Your email address (**joao.g.marques@tecnico.ulisboa.pt**) was recorded when you submitted this form.

Team number *

20

1st *

58722 João Granado Marques

2nd

76691 João Bernardo dos Santos Silva

3rd

84390 João Diogo Ferreira Bravo

Covid deaths analysis

There is a strong controversial about the real impact of COVID-19 on deaths around world. The present lab aims to study this in Portugal. And for that we have two data files available on Fénix (section labs): - the first one 'covid19_pt.csv' expresses the evolution of the number of deaths due to COVID-19 in Portugal; - the second one 'deaths_pt.csv' reports the number of deaths recorded in Portugal from 2015 to 2020. The lab has 4 sections. Please fill each one of them and submit the results.

Data Profiling

Provide a maximum of 5 image files (<=1M in size each) for profiling each datafile, and supporting the answers to the following questions .

Question 1 - Distribution

The best description for the distribution of values observed for...

	Uniform	Normal	Exponential	None of previous
covid19 is...	○	○	◉	○
deaths is...	○	○	○	◉

Question 2 - Stationarity

The series are stationary?

	Stationary	Non-stationary
covid19 is...	☑	☐
deaths is...	☑	☐

Question 3 - Granularity

The best granularity for...

	daily	weekly	monthly	quarterly	yearly	none of previous
covid19 is...	○	◉	○	○	○	○
deaths is...	○	◉	○	○	○	○

Question 4

Choose the True statements.

- [x] Both series have the same dimensionality.
- [ ] Both series are in the same granularity.
- [x] covid19 is an accumulated series, meaning its values are always increasing and each observation is the sum of all previous ones.
- [ ] The subtraction of one series to the other allows for discovering the number of additional deaths in 2020.

COVID-19 Profiling

Submitted files

- Q1 covid19 diff weekly variables distribution - João Granado Marques.png
- Q2 covid diff stationary 11 bins - João Granado Marques.png
- Q3 covid diff daily granularity - João Granado Marques.png
- Q3 covid diff weekly granularity - João Granado Marques.png
- Q4 covid original data - João Granado Marques.png

DEATHS_PT Profiling

Submitted files

- Q1 deaths weekly variables distribution - João Granado Marques.png

Q2 deaths stationary 12 bins - João Granado Marques.png

Q3 deaths monthly granularity - João Granado Marques.png

Q3 deaths weekly granularity - João Granado Marques.png

Data Transformation

Provide a maximum of 5 image files (<=1M in size each) for profiling each datafile, and supporting the answers to the following questions.

Question

Which are the most adequate operations to apply to each file for discovering the number of additional deaths in 2020.

	smoothing	aggregation (reducing granularity)	differencing	change of space	none
covid-19	☐	☑	☑	☐	☐
deaths	☑	☐	☐	☐	☐

COVID-19 Transformation

Submitted files

0 covid original - João Granado Marques.png

1 covid diff - João Granado Marques.png

2 covid diff weekly - João Granado Marques.png

DEATHS Transformation

Submitted files

0 deaths original - João Granado Marques.png

1 deaths smoothing over a monthly window - João Granado Marques.png

Forecasting

Predict the number of deaths in Portugal

DEATHS 2019

Train a predictor with ARIMA using all data from past years, to predict weekly deaths for 2019. Submit the chart showing the prediction.

Submitted files

DEATHS 2019 - João Granado Marques.png

DEATHS 2020

Train a predictor with ARIMA using all data from past years, to predict weekly deaths for 2020 (just until the first week of December).. Submit the chart showing the prediction.

Submitted files

DEATHS 2020 - João Granado Marques.png

DEATHS Dec 2020

Using the predictor trained in the previous task, predict the number of deaths for the remaining weeks of 2020, only using the data available in the deaths_pt file. Submit the chart showing the prediction.

Submitted files

DEATHS Dec 2020 - João Granado Marques.png

Evaluation

Report the RMSE for the predictors trained...

	< 1000	1000 <= x < 5000	5000 <= x < 10000	>= 10000
2019 predictor	◉	○	○	○
2020 predictor	◉	○	○	○

Motif Discovery

Using the original data from covid-19 file, find the 10 best motifs.

Motifs COVID-19

Submit one single chart showing the original data and each one of the motifs found.

Submitted files

covid motifs - João Granado Marques.png