**APPLIED STATISTICS EXAM**

**DATE:** 12/07/2022

**STUDENT:** FILIPPO CIPRIANI

**PERSONA CODE:** 10956877

**EXERCISE NUMBER 4**

Plotting the matplot of the daily values of traffic:

(If) Using a generalized cross-validation method we obtain that the optimal number of basis.a) che non vada ne a fare undersmoothing né oversmoothing

(if) plot of the basis:

Then we smooth the data using Fourier/Bspline basis with nbase elements and order 3.

if int conf:As in linear models, we can estimate the variance of x(t) as

# sigma^2\*diag[phi\*(phi'phi)^{-1}(phi)'], obtaining the interval:

To evaluate the first derivative from datas we use an incremental ratio approximation and for smooth function we use its first derivative. Then we can compare them:

pca:

Computing the pCA on our datas we obtain these first 3 coefficients obtained for Day 1:

screeplot:

variance explained along the first 3 functional principal components:

plot of the first 3 eigenfunctions.

Since the first 2 principal components explain the 90% of the variance so a possible dimensionality reduction can be done considering only these 2.

The first pc can be representative of…

Plotting the effect of the first 2 PCs on the mean we can affirm that ….

Using the k-mean alignment algorithm we can see that the datas are clustered according to

The first cluster …

**POINT A)**

**POINT B)**

**POINT C)**

**POINT D)**