UNIVERSITÉ DE MONTPELLIER

MASTER 1 BIOSTAT

Challenge2020/2021

Sidy Sow

2020 — 2021

Table des matières

0.1 Introduction

The goal of our project is to predict the number of vellos that will between 00H01 and 09H on April 2nd at the cyclcle track at the tram stop Albert streetcar stop first we will work on real data that are recorded which are recorded every day since March 2020 : : https ://docs.google.com/spreadsheets/d/e/2PACX-1vQVtdpXMHB4g9h75a0jw8Cs AIThodhVHNLxlZYm8fuoWj/p.

0.2 Prediction

After importing the data we saw that the first 2 lines were empty and rows were empty and NAN on the last two columns. columns. Therefore, in order to carry out our work we have deleted the last two rows and the first two rows and we have replace the NAN that the data by zeros and then we obtained a table with 3 columns table with 3 columns : in the first column we have DateTalent, the second the second one Bikes since January 1st / Grand total and the last Bikes today / Today last Bikes today / Today's total. What interests us more for this project is the last column of the table (Bikes today / Today's Today's total) our dataframe starts since March 12th. During this time, our prediction work consists then in working on the data between 00H01 and 9H.before starting our prediction we will delete the data of the day between 09H and 23H59. To carry out our work we will use the method of prediction ARIMA seen in course temporal series.

This method gives us the number of bicycles that pass every day between between 00H01 and 9H.This time there are very small values due to small values due to firewalls.
So the number of bikes expected on 02/April is 240.
Translated with www.DeepL.com/Translator (free version)