

Étude de cas:

Analyse de marché du transport aérien canadien avec R

Atelier d'introduction à R

BEAUCHEMIN, DAVID

CABRAL CRUZ, SAMUEL

GOULET, VINCENT

Dans le cadre du colloque R à Québec

25 mai 2017

Table des matières

| | |
|--|----------|
| Préface | 2 |
| Introduction | 4 |
| Étude de cas | 5 |
| 1.1 Extraction, traitement, visualisation et analyse des données . . . | 5 |
| 1.2 Création de fonctions utilitaires | 5 |
| 1.3 Communication des résultats | 5 |
| 1.4 Analyse de la compétition | 5 |
| 1.5 Ajustement de distribution statistiques sur données empiriques . | 5 |
| 1.6 Simulation et analyse de rentabilité | 5 |
| Conclusion | 6 |

Préface

Dans le cadre du colloque "R à Québec" qui se tiendra le 25 et 26 mai 2017 sur le campus de l'Université Laval, une séance d'introduction au langage de programmation R sera offerte aux participants. Cette séance vise principalement la compréhension et la pratique permettant de maîtriser les rudiments de cet environnement de programmation. [1] Cette séance sera divisée en deux parties. En ce qui concerne la première partie, les fondements du langage seront visités d'une manière théorique sous la forme d'un exposé magistral. La deuxième partie, tant qu'à elle, se concentrera davantage sur la mise en pratique des notions abordées lors de la première partie grâce à la complétion d'une étude de cas cherchant à faire l'analyse de marché du transport aérien canadien. Ce document correspond en fait à la documentation complète de cette deuxième partie de formation.

Étant donné qu'il s'agit tout de même d'une formation pour débutants, la majorité du code sera déjà fournie, mais il n'en vaut pas moins la peine de parcourir ce projet si ce n'est que pour constater la puissance et la simplicité du langage. De plus, il est souvent difficile de mettre en perspective les innombrables fonctionnalités d'un langage lorsque nous commençons à l'utiliser. Cet étude de cas nous fournit ainsi un bel exemple d'enchaînement de traitements jusqu'à l'aboutissement ultime qui consiste à répondre aux questions que nous nous posions avant même d'amorcer l'analyse.

D'autre part, il est important de préciser que le code qui sera présenté ne correspond pas toujours à la manière la plus efficiente d'accomplir une tâche donnée. L'objectif principal étant ici la transmission de connaissances dans un dessin éducatif plutôt que d'une réelle analyse de marché. Nous tenons aussi à mentionner que bien qu'il s'agit d'une formation s'adressant à des débutants, plusieurs notions qui seront mises en valeur font plutôt état de niveau intermédiaire et avancé, mais apportées toujours de manière simplifiée et accessible à quiconque qui n'aurait jamais travaillé avec R.

Nous tenons à remercier Vincent Goulet de nous avoir fait confiance dans l'élaboration de cette partie de la formation ainsi que tous les membres du comité organisationnel de l'événement. Nous croyons sincèrement que R est un langage d'actualité qui se révèle un atout à quiconque oeuvrant dans un domaine relié de près ou de loin aux mathématiques.

Introduction

Étude de cas

1.1 Extraction, traitement, visualisation et analyse des données

HELLO

1.2 Création de fonctions utilitaires

1.3 Communication des résultats

1.4 Analyse de la compétition

1.5 Ajustement de distribution statistiques sur données empiriques

1.6 Simulation et analyse de rentabilité

Conclusion

Bibliographie

[1] R à québec 2017.



setwd()

La fonction setwd permet de redéfinir la valeur du working directory

```
1 ##### Setting working directory properly #####
2 setwd('.')
3 (path <- getwd())
4 set.seed(31459)
5 ##### Question 1 – Extraction, traitement, visualisation et analyse
  des donnees #####
6 # 1.1 – Extraire les bases de donnees airports.dat et routes.dat
7 airports <- read.csv("https://raw.githubusercontent.com/jpatokal/
  openflights/master/data/airports.dat", header = FALSE,
  stringsAsFactors = TRUE, na.strings=c('\n',''))
8 routes <- read.csv("https://raw.githubusercontent.com/jpatokal/
  openflights/master/data/routes.dat", header = FALSE,
  stringsAsFactors = TRUE, na.strings=c('\n',''))
9 airlines <- read.csv("https://raw.githubusercontent.com/jpatokal/
  openflights/master/data/airlines.dat", header = FALSE,
  stringsAsFactors = TRUE, na.strings=c('\n',''))
```