

LSTAT2200

Lucien Ledune

mai 26, 2018, 13:52

Contents

Introduction	1
Jeu de données	1
Contraintes	2
Questions de l'étude	2
Quel type de poivron est le plus consommé à Legoland ?	2
Échantillonnage et calculs	2
Discussion	3
Est-ce que la consommation de poivron est homogène au sein des trois provinces de Legoland ?	3

Introduction

Dans le cadre du cours d'échantillonnage et sondage, nous allons réaliser un projet résumant la matière vue au cours. Ce travail sera divisé en plusieurs parties :

- Un descriptif de la problématique
- La description des méthodes de sondages utilisées
- Les statistiques descriptives de nos différents sondages
- Une discussion sur les résultats
- Une discussion sur les limites de notre approche et comment l'améliorer

Jeu de données

Pour ce travail nous utiliserons le jeu de données **Legoland** qui représente la consommation de poivrons des habitants de Legoland. Legoland se divise en trois provinces, et les principaux supermarchés vendant les poivrons sont eux aussi au nombre de trois. Les habitants se nourrissent de poivrons de trois différentes sortes : Rouge, jaune et vert.

Une ligne représente une transaction, que l'on peut associer à la consommation de poivrons d'un ménage au cours d'une journée.

Voici les différentes variables :

Province : La province dans laquelle la transaction a été effectuée.

Enseigne : Le supermarché dans lequel elle a été faite.

Paielement : Le type de paiement

Vert, jaune, rouge : La quantité de poivrons de chaque type achetée.

N_Personnes : Nombre de personnes au sein du ménage.

Il faut garder à l'esprit que même si nous avons la base de donnée pour ce projet, nous ne pouvons pas nous en servir directement puisque le but est d'échantillonner. Nous disposons cependant de certaines informations comme les effectifs des provinces et des enseignes.

Contraintes

- Chaque province de Legoland doit être représentée dans l'échantillon, ainsi que chaque enseigne dans chaque province.
- L'obtention des informations à un coût : 1, 1.2 et 1.5\$ pour Grancub, P'ti rond et Toupla.
- Le budget pour chaque échantillon est de 25000\$.

Questions de l'étude

Nous allons tenter de répondre à 4 questions principales au cours de ce travail :

- Quel type de poivron est le plus consommé à Legoland ?
- La consommation de poivrons est-elle homogène au sein des trois provinces ?
- Le chiffre d'affaire diffère-t-il au sein des chaînes de supermarchés ?
- Est-ce que l'installation de terminaux pour le paiement électronique serait rentable ?

Quel type de poivron est le plus consommé à Legoland ?

Échantillonnage et calculs

Puisque dans cette question nous ne nous intéressons pas aux différents groupes mais bien à la totalité de la population, un sondage aléatoire sans remise devrait être efficace.

Premièrement nous allons constituer notre échantillon. Celui-ci sera fait grâce à la fonction *srswor* afin de prendre 19250 échantillons.

Nous allons ensuite vérifier si celui-ci ne dépasse pas la limite de coûts, auquel cas nous devrions revoir ce chiffre à la baisse.

[1] 24913.14

Nous sommes juste en dessous de la limite, nous pouvons donc conserver cet échantillon pour l'analyse.

Nous allons maintenant vérifier si cet échantillon représente bien les trois enseignes et les trois provinces. Pour ce faire nous allons afficher calculer les tables de proportions des provinces ainsi que des enseignes de la population (dont nous connaissons les totaux à priori) et les comparer avec les proportions de notre échantillon. Nous pourrions ensuite afficher la différence de ces proportions entre l'échantillon et l'information à priori pour voir si nous en sommes proche.

Nord	Centre	Sud
0.0004559	-0.001353	0.0008971

Grancub	P'ti rond	Toupla
0.00454	-0.006231	0.001691

En effet il semble que les proportions soient bien respectées. Nous pouvons également afficher la table croisée de notre échantillon :

	Grancub	P'ti rond	Toupla
Nord	2150	2141	2176
Centre	2183	2065	2112
Sud	2169	2118	2136

L'échantillon étant maintenant constitué et validé nous pouvons maintenant répondre à la question en faisant la somme des achats des différents types de poivrons de l'échantillon :

vert	jaune	rouge
161140	163636	333455

Discussion

Comme nous le voyons très bien ici, les poivrons rouges sont largement majoritaires avec plus de 333.000 poivrons consommés. Les deux autres types de poivrons sont assez similaire dans leurs ventes avec 163636 et 161140 en faveur des poivrons jaunes. Cette information pourrait être intéressante pour la détermination des prix des poivrons, car elle nous informe sur la demande de ceux-ci. Cependant avant de prendre une conclusion trop hâtive nous allons aller plus loin en regardant si cette consommation est homogène entre les différentes provinces.

Est-ce que la consommation de poivron est homogène au sein des trois provinces de Legoland ?