

Data Visualisation avec Python à partir des fichiers csv

Réalisée par
ABBA Dada

Introduction

le langage python prend une ampleur tout à fait inattendue avec une utilisation de plus en plus large. En science des données, python est devenu le langage de prédilection pour le traitement et l'analyse des données. Python en data science est devenu une référence.

Python est le plus utilisé en data science grâce à 3 facteurs :

1. *La simplicité du langage* : pour un langage de programmation orienté objet, python possède une courbe d'apprentissage ultra-rapide. Vous êtes très rapidement opérationnel en python. Quelques jours suffisent à l'acquisition des bases du langage et à vous rendre opérationnel.
2. *La multitude de bibliothèques* (librairies ou packages suivant la terminologie) : mettre en place une bibliothèque en python est extrêmement simple et ceci a permis la publication de bibliothèques spécialisées par des équipes de recherche.
3. *Le nombre impressionnant d'API* vers d'autres programmes ou d'autres environnements. Il est extrêmement simple de se connecter à d'autres environnement avec python.

Objectifs

1. Créez une interface qui permet d'afficher des données à partir des fichiers csv.
2. Tracer des courbes.
3. Afficher le maximum, minimum, moyenne, écart-type, somme.

Outils utilisés

1) Tkinter(gui)

Tkinter c'est un module présent par défaut dans Python.

Ce module permet de créer des interfaces graphiques en offrant une passerelle entre Python et la bibliothèque Tk.

2) Pandas

Avec Numpy et Matplotlib , la librairie Pandas fait partie des librairies de base pour la data science en Python. Pandas fournit des structures de données puissantes et simples à utiliser, ainsi que les moyens d'opérer rapidement des opérations sur ces structures.

Algorithme

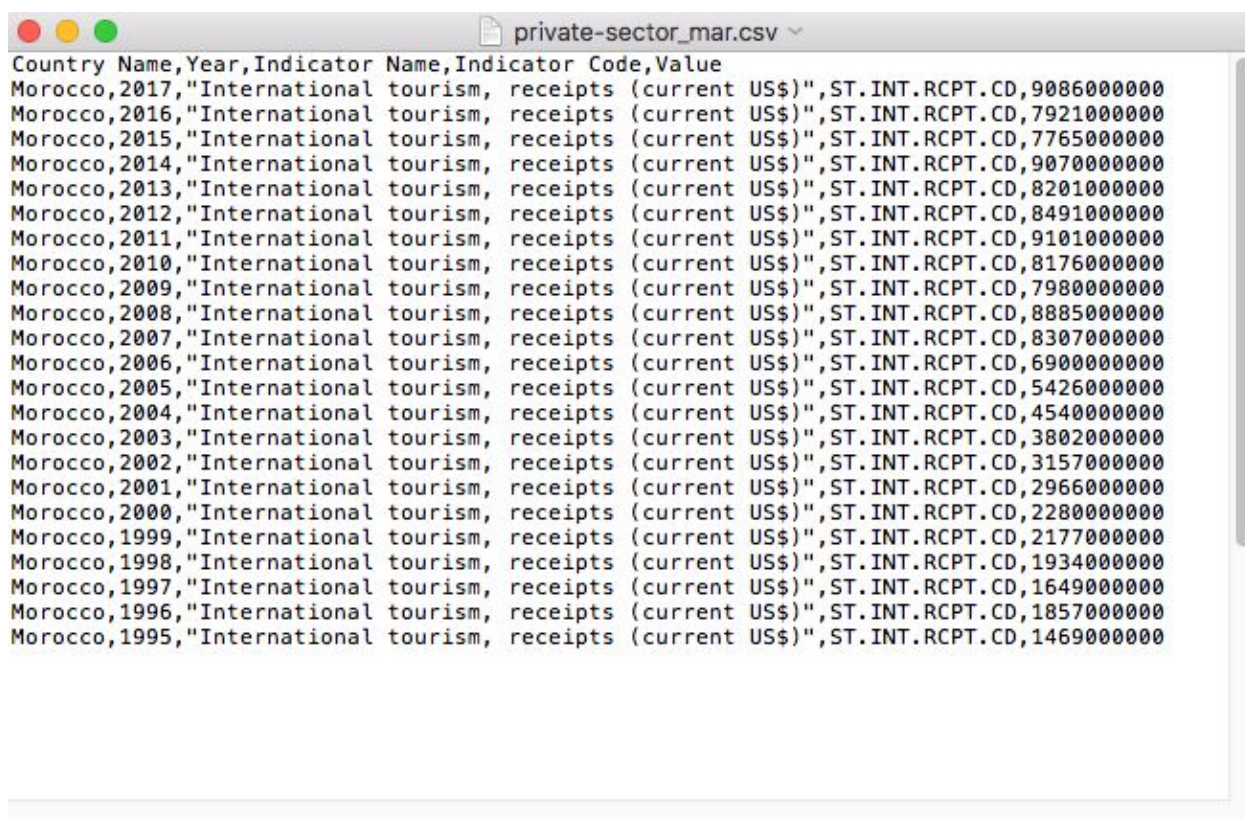
1. Lire les données du fichier csv en utilisant fonction `read_csv()` de package pandas.
2. Organisez ces données sous forme de tableau via Treeview de package ttk
3. Affichez à l'utilisateur les variables quantitative du fichier passé ,pour afficher sa min,max,mean...
4. Après que l'utilisateur confirme son choix affichez les statistique et les graphes.

Les données

C'est difficile de trouver des données complètes sur Maroc, mais j'étais capable de trouver des fichiers csv sur site web :

<https://data.humdata.org/> fournit par organisation internationale OCHA.

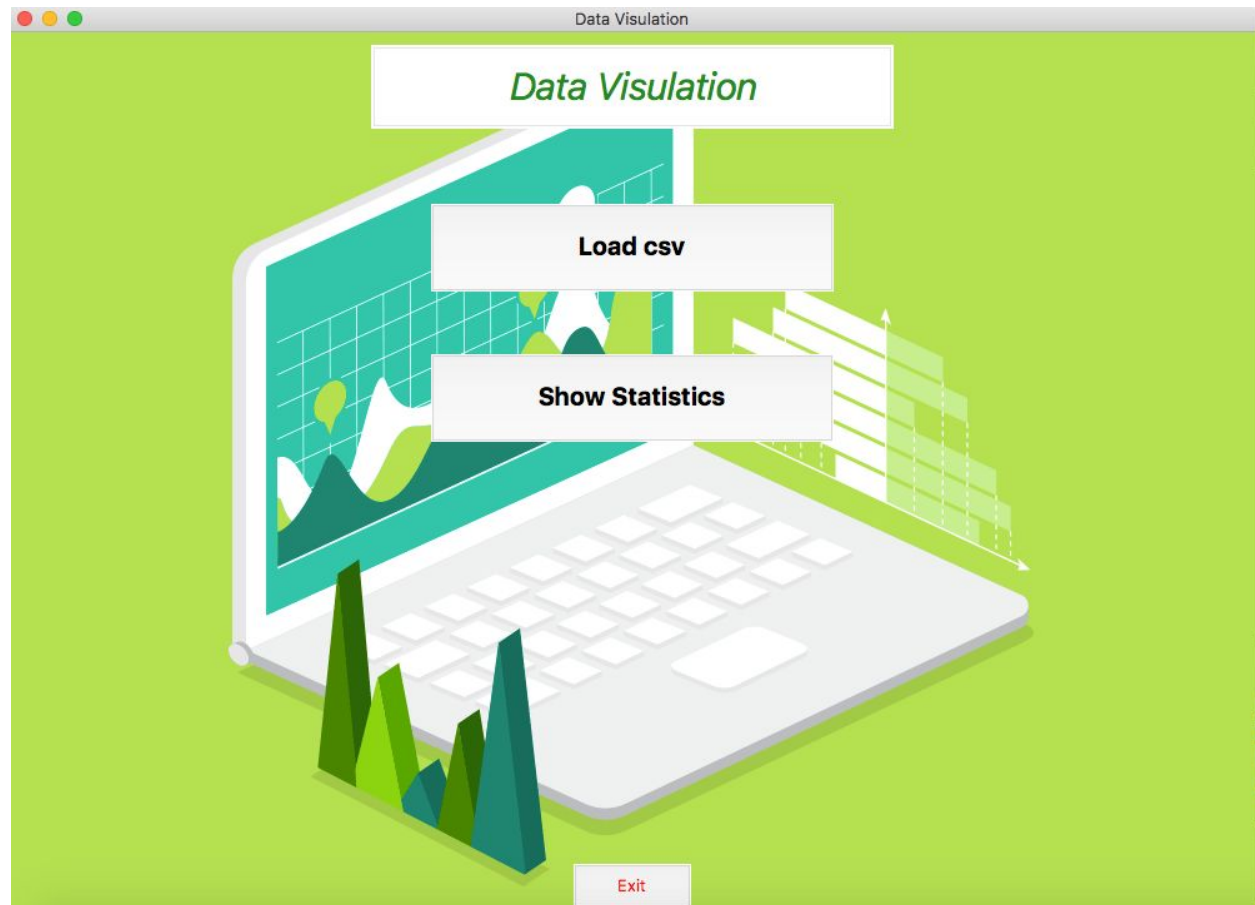
J'ai choisie le fichier qui concerne secteur privé plus précisément le revenu de la tourisme internationale .



Country Name	Year	Indicator Name	Indicator Code	Value
Morocco	2017	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	9086000000
Morocco	2016	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	7921000000
Morocco	2015	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	7765000000
Morocco	2014	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	9070000000
Morocco	2013	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	8201000000
Morocco	2012	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	8491000000
Morocco	2011	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	9101000000
Morocco	2010	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	8176000000
Morocco	2009	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	7980000000
Morocco	2008	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	8885000000
Morocco	2007	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	8307000000
Morocco	2006	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	6900000000
Morocco	2005	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	5426000000
Morocco	2004	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	4540000000
Morocco	2003	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	3802000000
Morocco	2002	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	3157000000
Morocco	2001	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	2966000000
Morocco	2000	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	2280000000
Morocco	1999	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	2177000000
Morocco	1998	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	1934000000
Morocco	1997	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	1649000000
Morocco	1996	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	1857000000
Morocco	1995	"International tourism, receipts (current US\$)"	ST.INT.RCPT.CD	1469000000

L'application(ScreenShot)

1. Homepage:



2. Load csv:

	Country Name	Year	Indicator Name	Indicator Code	Value
item0	Morocco	2017	International tourism, receipts	ST.INT.RCPT.CD	9086000000
item1	Morocco	2016	International tourism, receipts	ST.INT.RCPT.CD	7921000000
item2	Morocco	2015	International tourism, receipts	ST.INT.RCPT.CD	7765000000
item3	Morocco	2014	International tourism, receipts	ST.INT.RCPT.CD	9070000000
item4	Morocco	2013	International tourism, receipts	ST.INT.RCPT.CD	8201000000
item5	Morocco	2012	International tourism, receipts	ST.INT.RCPT.CD	8491000000
item6	Morocco	2011	International tourism, receipts	ST.INT.RCPT.CD	9101000000
item7	Morocco	2010	International tourism, receipts	ST.INT.RCPT.CD	8176000000
item8	Morocco	2009	International tourism, receipts	ST.INT.RCPT.CD	7980000000
item9	Morocco	2008	International tourism, receipts	ST.INT.RCPT.CD	8885000000

3. Show Statistics:

la fenêtre "window" nous demande de choisir une variable quantitative sur laquelle on va appliquer statistique descriptive.

