C2 - Python for Data Science

C-DAT-100

Dataframe and diagrams

Let's plot...to take over the world ?

1.0

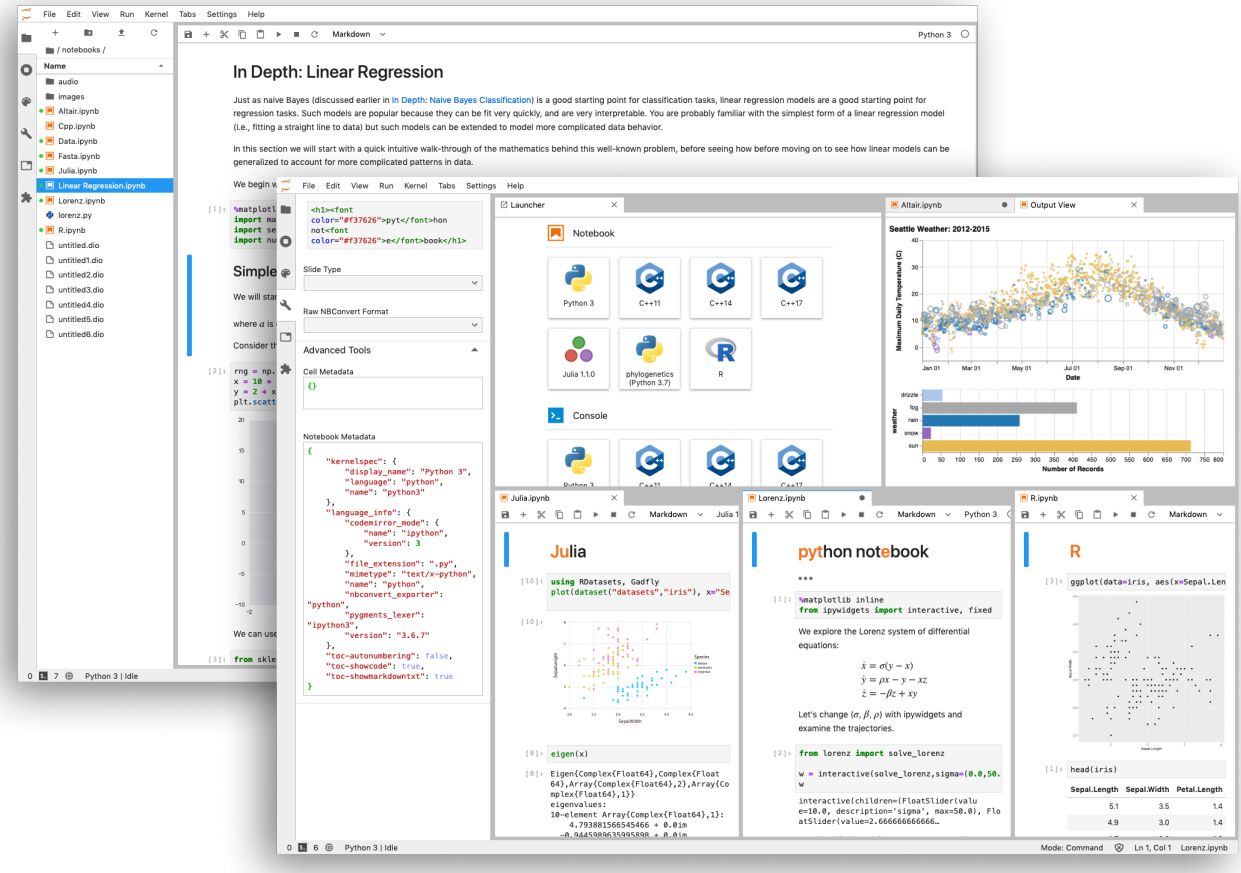
Dataframe and diagram

delivery method: py02 on Github
language: python

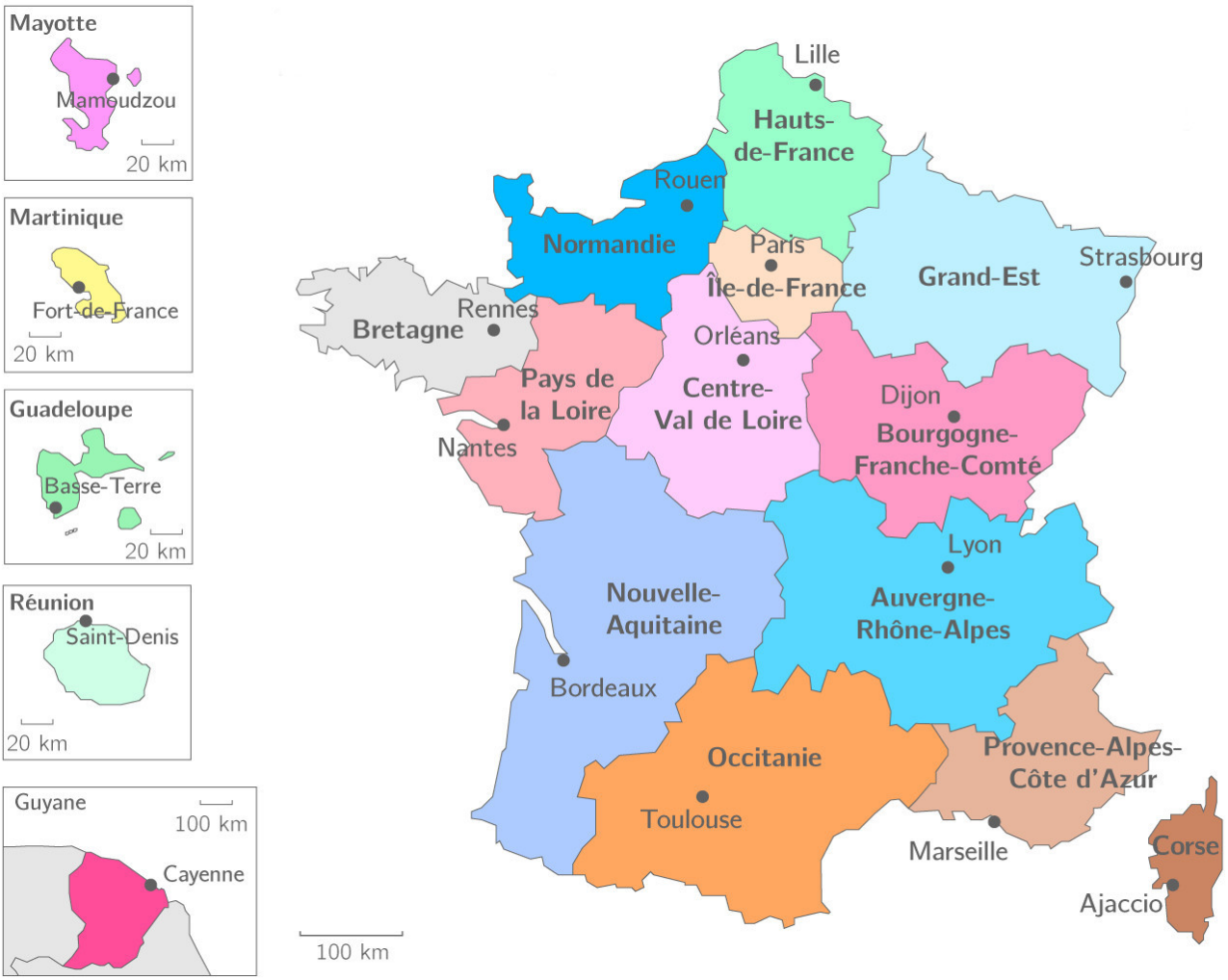
From now on, you must use **Jupyter Notebook** to code. It is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text.

You must only turn-in a single Notebook (.ipynb) containing all exercices.

You can use VS Code extention for Jupyter Notebook.



YOUR FIRST DATASET



This dataset is open source data from regions of France. Its contains all their income and public speding acording year by year from 2012 to 2019.

Download here : `https://epitechfr.sharepoint.com/:f:/s/Peda/EoopUAZBOWtBnd8LWM1Zo-4Bjyrzb6GOOONQdg7BTIbwfQ?e=vanJuj`

year	reg_name	tax_name	amount
2012	Bretagne	Charges financières	8452077.24
2015	Guyane	Autres dépenses d'investissement	8454368.16
2015	La Réunion	Charges financières	8455398.75
2015	La Réunion	Autres dépenses de fonctionnement	8466607.05
2014	Provence-Alpes-Côte d'Azur	Ventes de biens et services	8499274.13
2016	Île-de-France	Produit des cesssions d'immobilisations	8507776.74
2013	Bourgogne-Franche-Comté	Capacité ou besoin de financement	8512891.52
2013	Centre-Val de Loire	Autres dépenses d'investissement	8514274.36
2018	Guadeloupe	FCTVA	8536022.49
2019	Centre-Val de Loire	Charges financières	8540840.62
2014	Grand Est	FCTVA	8544653.8

EXERCISE 01 (5PT)

Display income of “Auvergne-Rhône-Alpes” county from the CVAE enterprise tax for all year (if there is more than one CVAE income for by year, you must sum all CVAE income from this year)

You must use **Pandas dataframe**.

EXEMPLE

```
Auvergne-Rhone-Alpes CVAE tax :
    2012 -> X euros.
    2013 -> Y euros.
    ...
```

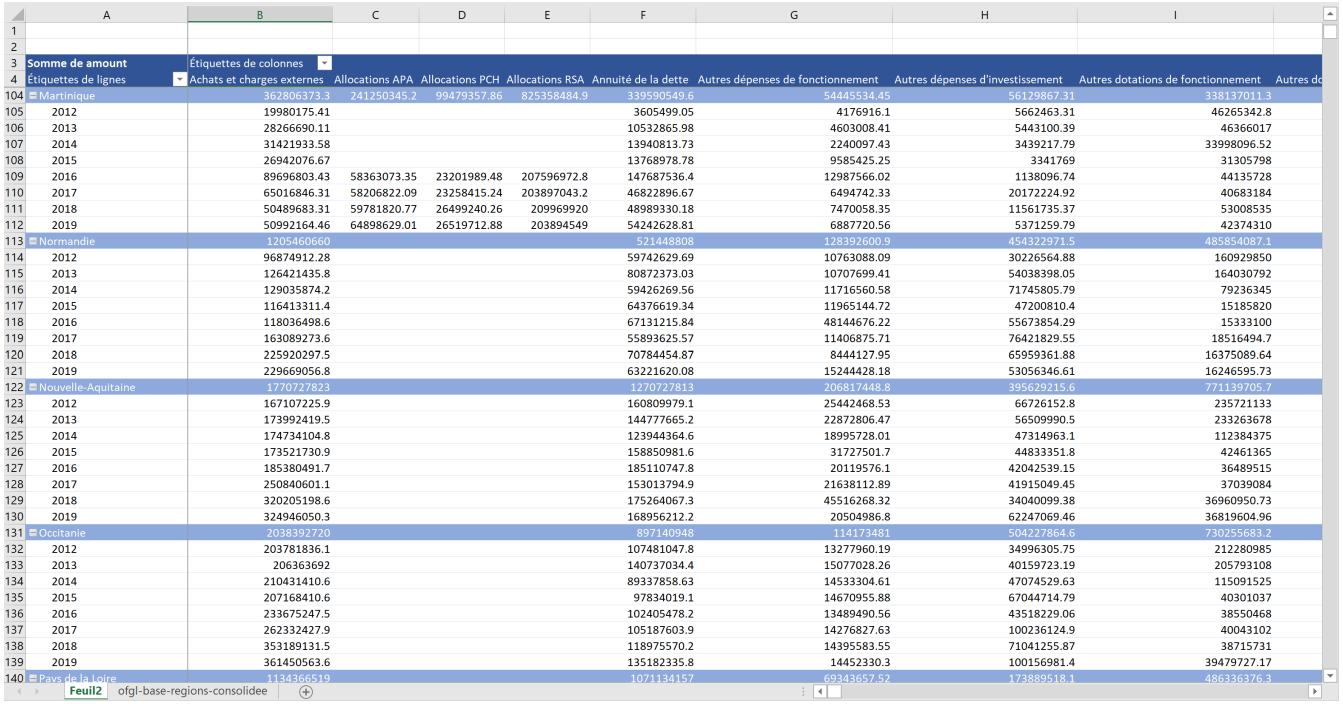
https://pandas.pydata.org/docs/getting_started/index.html

Exercise 02 (5pt)

Do the same as exercice 1 but for all regions of France sorted by taxes for all years.

As a exemple, this is a famous **tableau croisé dynamique** (pivot table) done by Excel

Example done with Microsoft Excel



Somme de amount	Étiquettes de colonnes								
Étiquettes de lignes	Achats et charges externes	Allocations APA	Allocations PCH	Allocations RSA	Annuité de la dette	Autres dépenses de fonctionnement	Autres dépenses d'investissement	Autres dotations de fonctionnement	Autres do
Martinique	362806373.3	241250345.2	99479357.86	825358484.9	339590549.6	54445534.45	56129867.31	338137011.3	
2012	19980175.41				3605499.05	4176916.1	5662463.31	46265342.8	
2013	28266690.11				10532865.98	4603008.41	5443100.39	46366017	
2014	31421933.58				13940813.73	2240097.43	3439217.79	33998096.52	
2015	26942076.67				13768978.78	9585425.25	3341769	31305798	
2016	89696803.43	58363073.35	23201989.48	207596972.8	147687536.4	12987566.02	1138096.74	44135728	
2017	65016846.31	58206822.09	23258415.24	203897043.2	46822896.67	6494742.33	20172224.92	40683184	
2018	50489683.31	59781820.77	26499240.26	209969920	48989330.18	7470058.35	11561735.37	53008535	
2019	50992164.46	64898629.01	26519712.88	203894549	54242628.81	6887720.56	5371259.79	42374310	
Normandie	1205460660				521448808	128392600.9	454322971.5	485854087.1	
2012	96874912.28				59742629.69	10763088.09	30226564.88	160929850	
2013	126421435.8				80872373.03	10707699.41	54038398.05	164030792	
2014	129035874.2				59426269.56	11716560.58	71745805.79	79236345	
2015	116413311.4				64376619.34	11965144.72	47200810.4	15185820	
2016	118036498.6				67131215.84	48144676.22	55673854.29	15333100	
2017	163089273.6				55893625.57	11406875.71	76421829.55	18516494.7	
2018	225920297.5				70784454.87	8444127.95	65959361.88	16375089.64	
2019	229669056.8				63221620.08	15244428.18	53056346.61	16246595.73	
Nouvelle-Aquitaine	1770727823				1270727813	206817448.8	395629215.6	771139705.7	
2012	167107225.9				160809979.1	25442468.53	66726152.8	235721133	
2013	173992419.5				144777665.2	22872806.47	56509990.5	233263678	
2014	174734104.8				123944364.6	18995728.01	47314963.1	112384375	
2015	173521730.9				158850981.6	31727501.7	44833351.8	42461365	
2016	185380491.7				185110747.8	20119576.1	42042539.15	36489515	
2017	250840601.1				153013794.9	21638112.89	41915049.45	37039084	
2018	320205198.6				175264067.3	45516268.32	34040099.38	36960950.73	
2019	324946050.3				168956212.2	20504986.8	62247069.46	36819604.96	
Occitanie	2038392720				897140948	114173481	504227864.6	730255683.2	
2012	203781836.1				107481047.8	13277960.19	34996305.75	212280985	
2013	206363692				140737034.4	15077028.26	40159723.19	205793108	
2014	210431410.6				89337858.63	14533304.61	47074529.63	115091525	
2015	207168410.6				97834019.1	14670955.88	67044714.79	40301037	
2016	233675247.5				102405478.2	13489490.56	43518229.06	38550468	
2017	262332427.9				105187603.9	14276827.63	100236124.9	40043102	
2018	353189131.5				118975570.2	14395583.55	71041255.87	38715731	
2019	361450563.6				135182335.8	14452330.3	100156981.4	39479727.17	
Pays de la Loire	1134366519				1071134157	69343657.52	173889518.1	486336376.3	

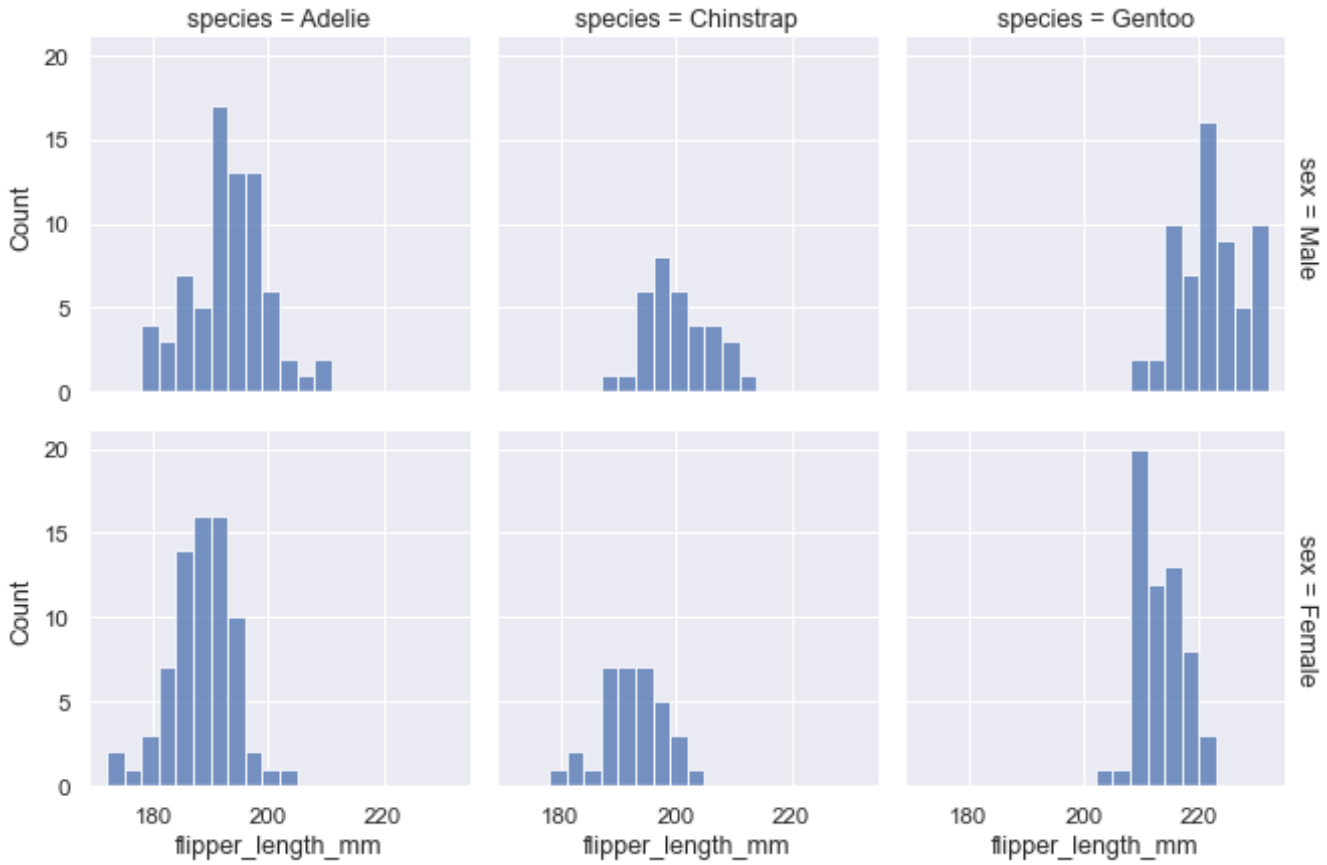
Feuil2 ofgl-base-regions-consolidee

Exercise 03 (5pt)

Make a chart to plot exercice 1.

You must use **Seaborn** python library : https://seaborn.pydata.org/tutorial/function_overview.html

Non-related exemple



this is Pinguin species is you want to know

EXERCISE 04 (5PT)

Make a chart the pivot table from execice 2.

As plotting a chart with 3 differents parameters is impossible, find the best practical solution.