



UTIU
UNIVERSITÀ TELEMATICA
INTERNAZIONALE **UNINETTUNO**

Corso di ingegneria informatica

A.A. 2020/2021

Data Visualization con Plotly

Analisi e visualizzazione dati occupazionali Italia (1996-2018) (ISTAT)

Nome dello studente:	<i>Antonello Cesare</i>
Matricola	<i>347HHHINGINFOR</i>
Data Appello	<i>Maggio 2021</i>
Esame	<i>Big Data Analysis & Visualization</i>

Indice

Sommario	3
Anaconda Navigator.....	4
Installazione su ambienti Linux e accesso a Jupyter	4
Esecuzione del Notebook Jupyter	5
Preparazione	5
Prima esecuzione e installazione dei moduli.....	6
Consultazione.....	6
Chiusura delle visualizzazioni Dash.....	8
Arresto dei Notebook Jupyter	8

Sommario

Il presente documento illustra i passaggi per la predisposizione dell'ambiente Anaconda basato su Python 3.8 per l'esecuzione del Notebook Jupyter *“Analisi e visualizzazione dei dati del Lavoro in Italia (1996-2018) (ISTAT)”*, progetto di lavoro del corso di *Big Data Analysis & Visualization*.

Vengono descritti i semplici passaggi per l'installazione dell'ambiente Anaconda Navigator su Linux, tuttavia l'installazione su Windows risulta essere assolutamente analoga.

Sono quindi descritti i passaggi per la predisposizione dell'ambiente Jupyter e vengono riportati alcuni screenshot che si potranno navigare in fase di esecuzione.

Anaconda Navigator

Installazione su ambienti Linux e accesso a Jupyter

Le istruzioni di installazione dell'ambiente Anaconda Navigator sono reperibili al seguente link: <https://docs.anaconda.com/anaconda/install/>.

Il pacchetto di installazione per tutti i SO è disponibile al seguente link: <https://docs.anaconda.com/anaconda/install/>, in ambiente linux viene scaricato un "file bash" .sh che può essere lanciato da console, è importante seguire le istruzioni e verificare dipendenze e prerequisiti indicati nella pagina di istruzioni <https://docs.anaconda.com/anaconda/install/linux/>.

In particolare, occorre verificare le dipendenze installando i pacchetti:

- `sudo apt-get install libgl1-mesa-glx libegl1-mesa libxrandr2 libxrandr2 libxss1 libxcursor1 libxcomposite1 libasound2 libxi6 libxtst6`

Si procede quindi lanciando il bash-script:

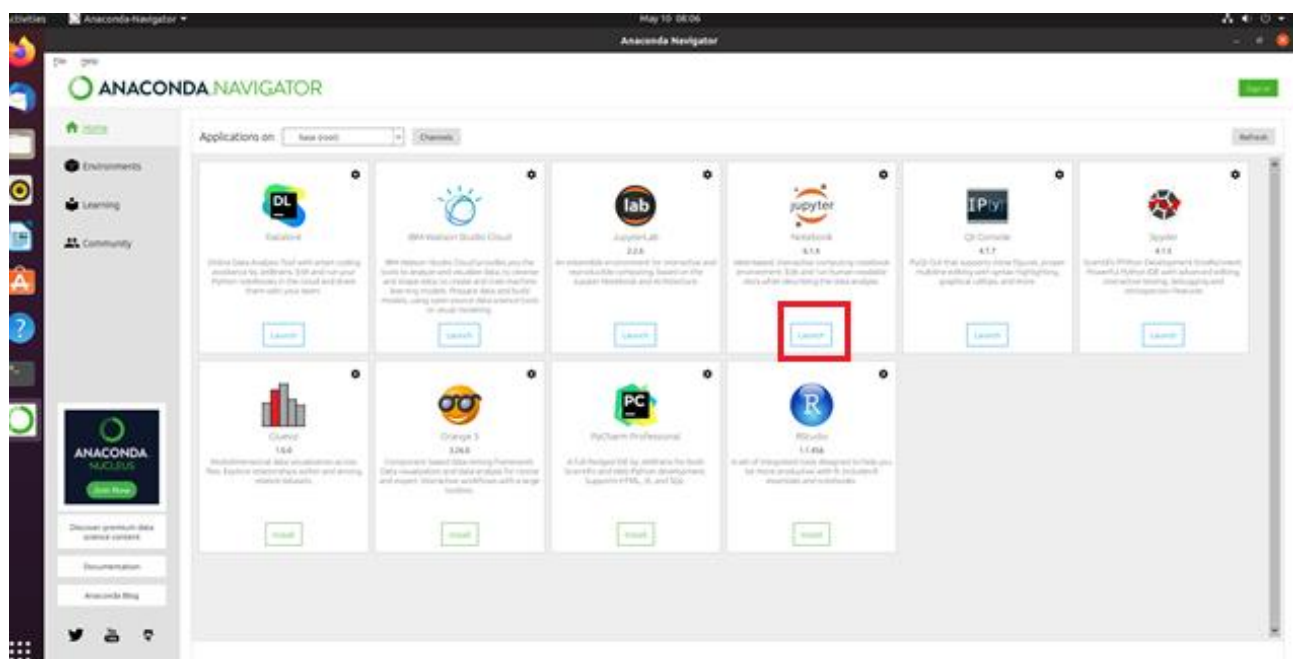
- `bash ./Downloads/Anaconda3-2020.11-Linux-x86_64.sh`

Per default il software viene installato nella home dell'utente all'interno del folder "anaconda3", al termine dell'installazione la procedura richiede se procedere all'inizializzazione, è consigliabile rispondere "yes".

Se tutto è andato a buon fine è possibile lanciare il navigator:

- `anaconda-navigator`

Una volta installato l'ambiente Anaconda, è possibile lanciare l'ambiente Jupyter:



Esecuzione del Notebook Jupyter

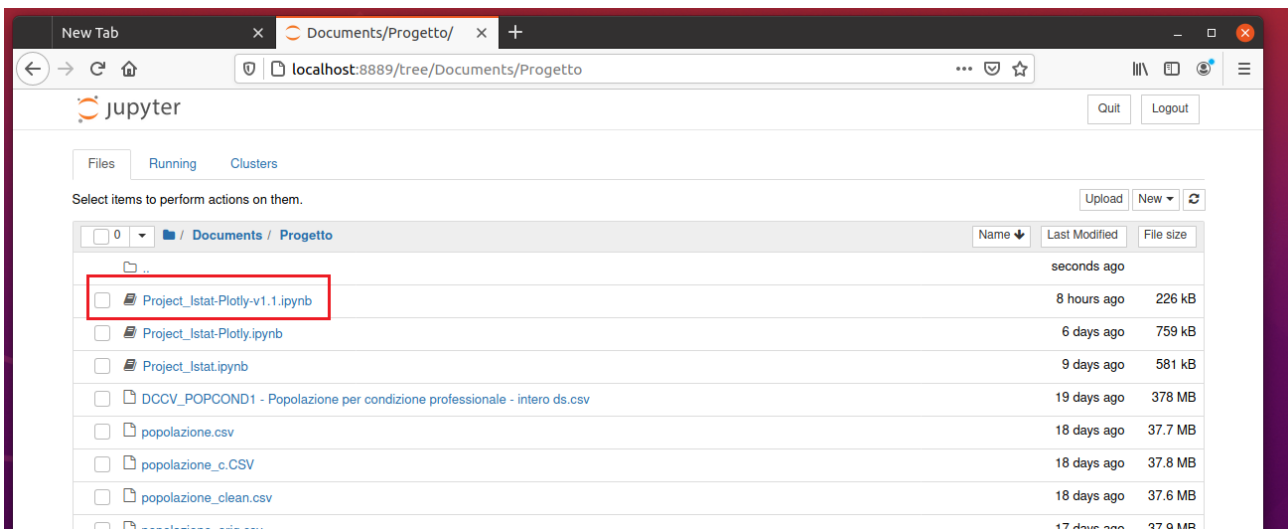
Preparazione

L'ultima versione del notebook e il dataset utilizzato sono scaricabili dal progetto GitHub:

https://github.com/caesar73-git/BDA_VIZ.git

È consigliabile scaricarlo localmente e decomprimere il dataset .csv nella stessa directory.

Una volta lanciato Jupyter è possibile navigare all'interno dell'interfaccia web (<http://localhost:8889/tree>) per cercare il file “.ipynb”.



Cliccando sul file corrispondente viene avviato il notebook.



Prima esecuzione e installazione dei moduli

I passi di installazione sono illustrati all'interno della pagina ufficiale [Getting Started with Plotly | Python | Plotly](#), possono essere lanciati anche direttamente dal notebook, ma può essere comodo lanciarli da Bash Terminal, sia attraverso comando `pip` che attraverso `conda`, riporto la procedura:

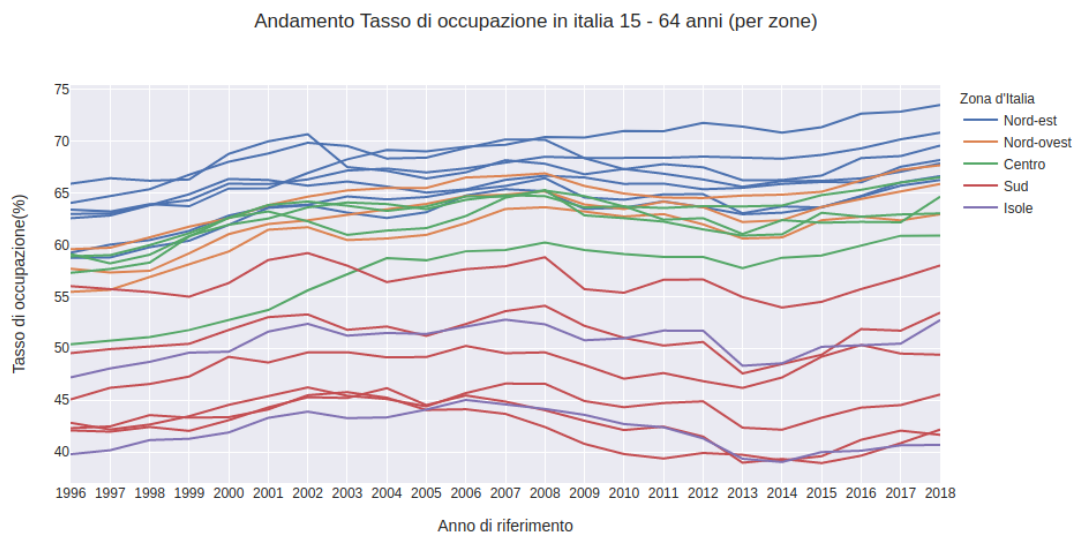
```
1. conda install -c plotly plotly=4.14.3
```

Installare il supporto Dash:

```
2. conda install -c conda-forge dash
```

Consultazione

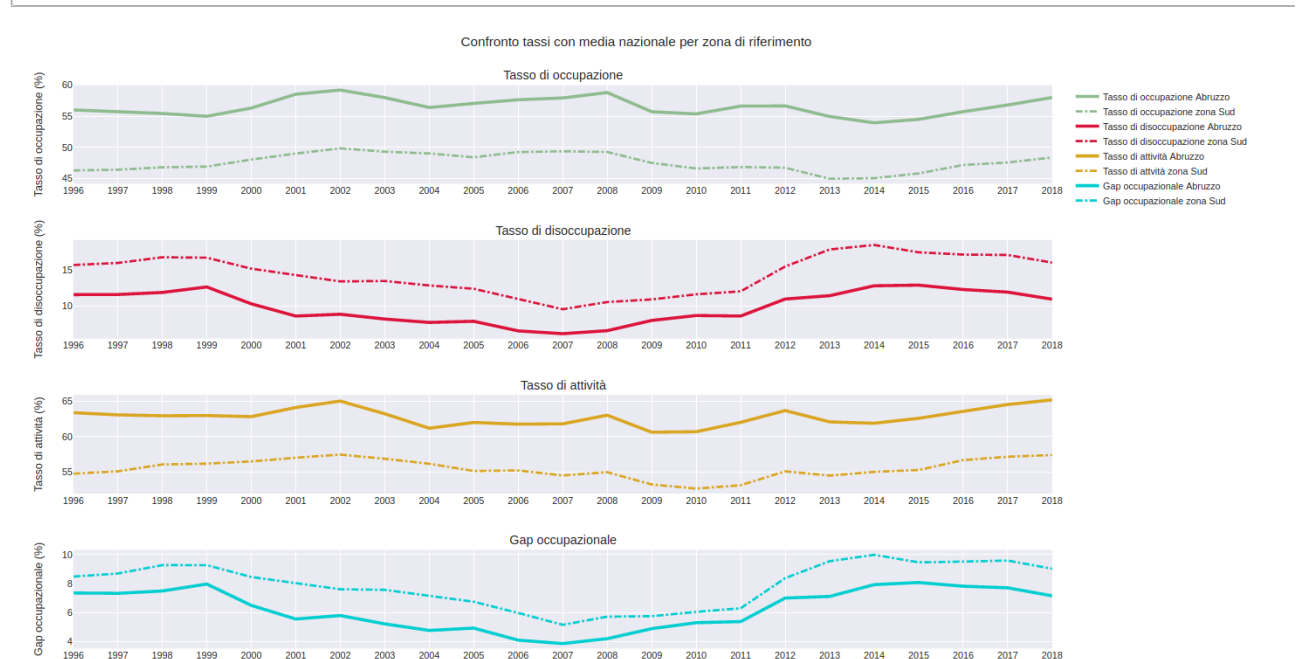
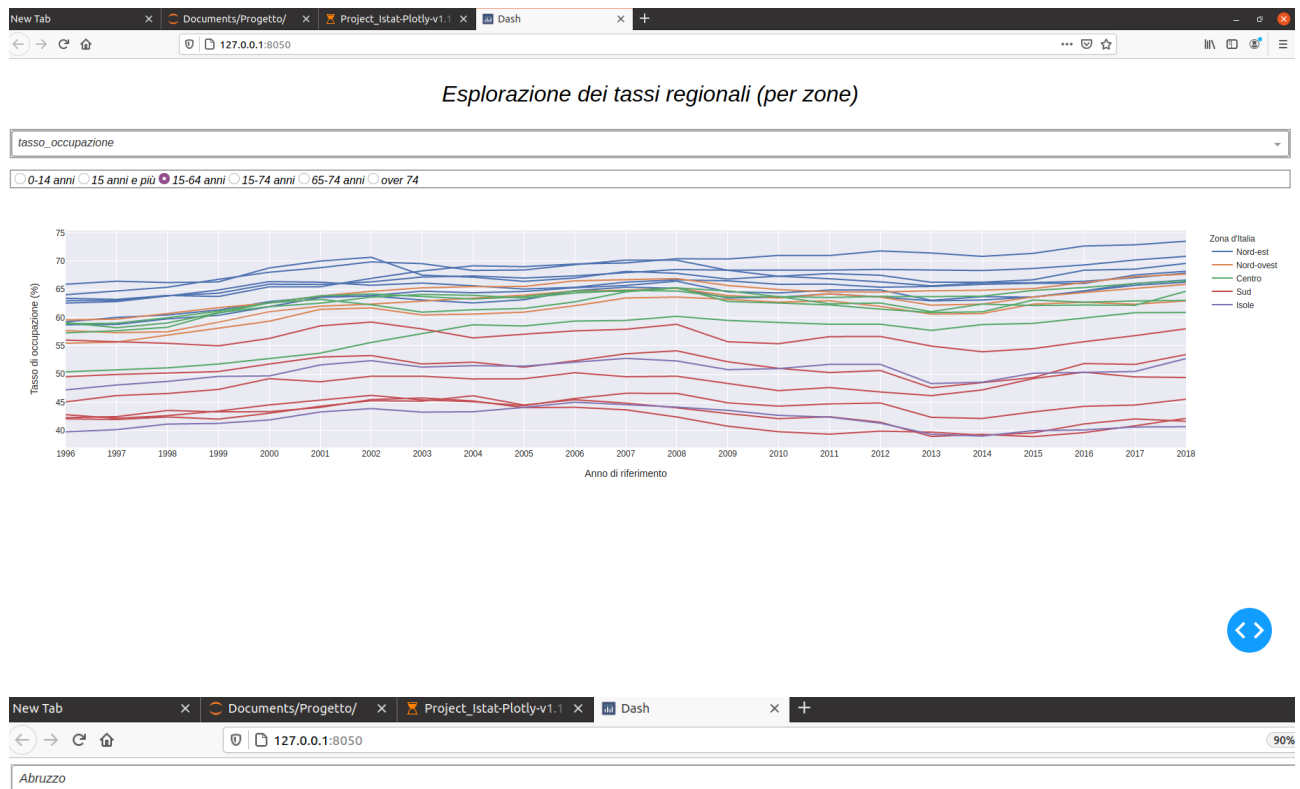
A questo punto l'esecuzione del notebook è quella standard, la visualizzazione dei grafici "standard" con Plotly e Plotly Express viene visualizzata *inline*

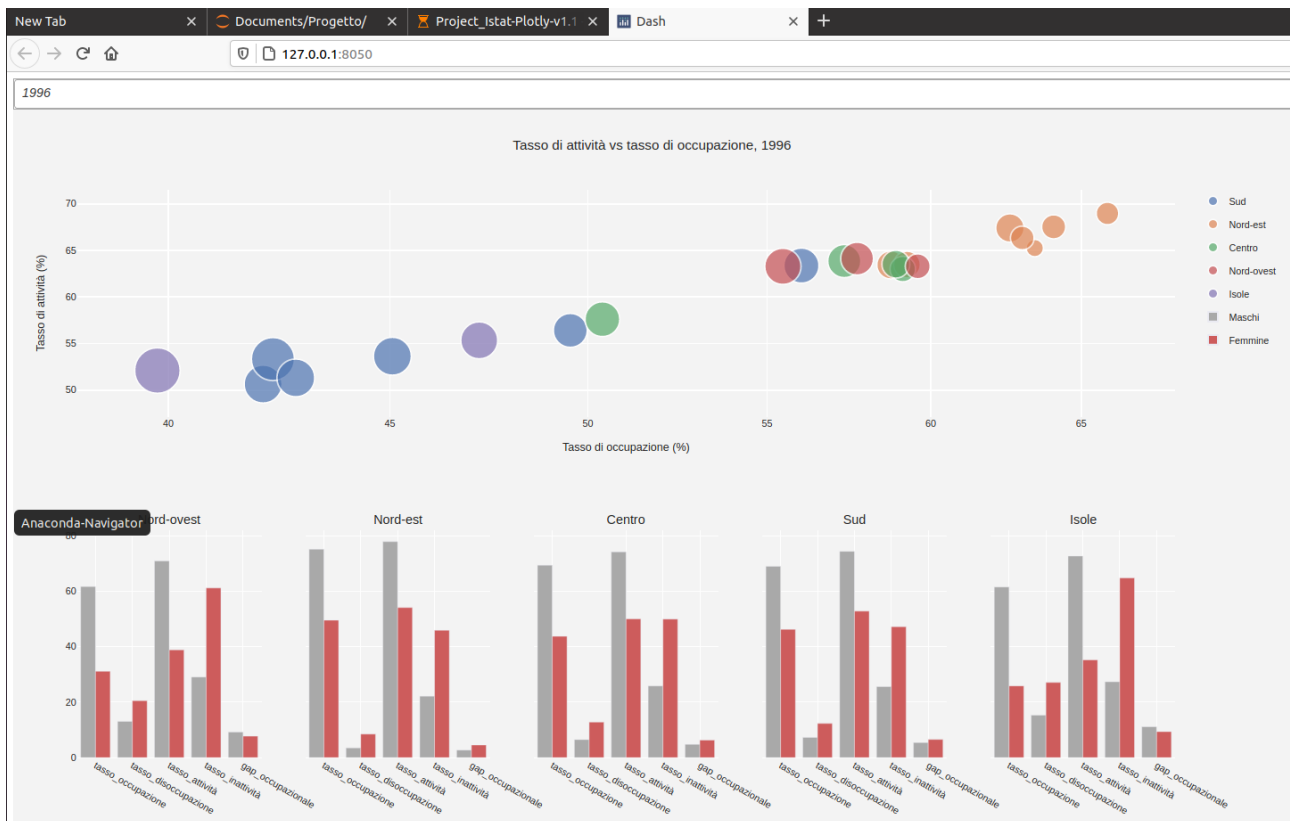


Esplorazione completa dei tassi

Analogo al grafico precedente ma consente di navigare tra i diversi tassi e scegliere la classe di età d'interesse. Oltre alla conferma che la classe di età "standard" è quella tra i 14 e 64 anni, si possono apprezzare alcune particolarità regionali sul tasso di attività/inattività anche nelle fasce 65-74. Uno degli elementi più significativi riguarda l'andamento della disoccupazione che tra il 2006 e il 2011 ha raggiunto valori minimi, pertanto è interessante

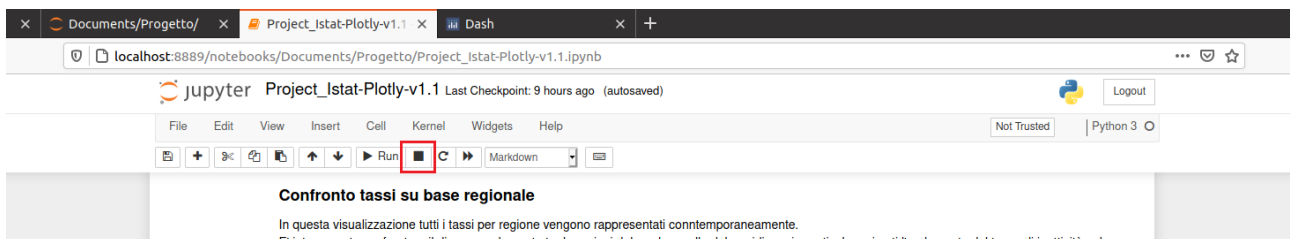
Le visualizzazioni create con Plotly Dash vengno aperte in un nuovo *tab* del browser:





Chiusura delle visualizzazioni Dash

Come mostrato le visualizzazioni Dash vengono eseguite su un web server locale che risponde alla porta 8050 (<http://127.0.0.1:8050/>), per arrestare l'esecuzione della dashboard è necessario interrompere dalla toolbar di Jupyter, diversamente non sarà possibile eseguire celle di codice successive.



Arresto dei Notebook Jupyter

Il monitoraggio e l'arresto dei notebook è gestito dalla Home Page di Jupyter:

