

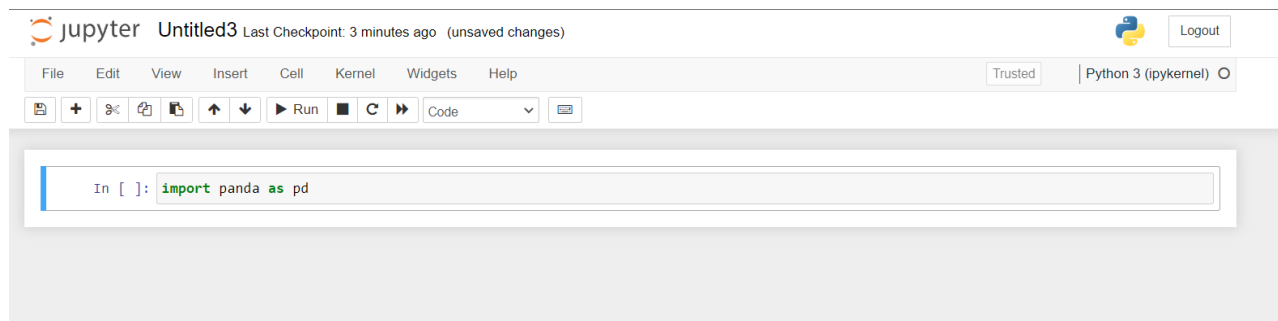
# Sprint 1 - Dataset exploration

We have used Jupyter notebook to explore the dataset.

## 1. Creating a new notebook



## 2. Loading the data into jupyter notebook by using a powerful python library called Pandas (for Python Data Analysis Library).



### 3. Understanding the dataset

data = pd.read\_csv("Global\_Superstore2.csv", encoding = 'unicode\_escape', engine = 'python')

data

Python

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...	Product ID	Category	Sub-Category	Product Name	Sales
0	32298	CA-124891	2012-07-2012	31-07-2012	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...	TEC-AC-10003033	Technology	Accessories	Plantronics CS510 - Over-the-Head monaural Wir...	2309.650
1	26341	IN-2013-77878	2013-02-2013	05-02-2013	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...	FUR-CH-10003950	Furniture	Chairs	Novimex Executive Leather Armchair, Black	3709.395
2	25330	IN-2013-71249	2013-10-2013	17-10-2013	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...	TEC-PH-10004664	Technology	Phones	Nokia Smart Phone, with Caller ID	5175.171
3	13524	ES-2013-	28-01-	30-01-	First Class	KM-16376	Katherine Murray	Home Office	Berlin	Berlin	...	TEC-PH-10004664	Technology	Phones	Motorola Smart Phone	2892.510