**Abstract :**

Amazon start his virtual doors in 1995 sold only his own products. In 2005, Bezos the founder and CEO ,decide to open the website virtual door to vendors. And he said “If a third party could offer a better price or better availability on a particular item, then we wanted our customer to get easy access to that offer.”

In this dataset (downloaded from ProPublica) we will go in deep with **Amazon’s shopping algorithm.**

Our dataset listings for 250 bestselling products across a wide range of categories, from electronics to household supplies, over a period of several weeks during summer 2016

We compared pricing and shipping costs for products offered by multiple vendors, including those sold by Amazon and sellers in the "Fulfilled by Amazon" program.

***Definition:***

* **Buy Box** = Items were selected for the most prominent placement on Amazon’s virtual shelves that pops up first as a suggested purchase.

One of a shareholder asked about Amazon’s practice of promoting products sold by other companies on its website. Bezos replied that the company had “very objective customer-centered algorithms‘’ that automatically award the “buy box” to the lowest price seller, provided “they actually have it in stock and can deliver it.”

So Buy box should rank items lowest price seller.

* **Prime program** = aiming customer with annul subscription gives free shipping and many features.
* **Fulfilled by Amazon FBA** = aiming vendors with annul subscription

Ranking Items In buy box is real? Is all product have same Procedure?

Example to explain the facts:

Product A

vendor 1 = $6.75 with free shipping

vendor 2= $7.27 with free shipping.

The computer program brushed aside those offers, instead selecting **Product A** sold by Amazon itself for slightly more, $7.80. This seemed like a plausible choice until another click of the mouse revealed shipping costs of $6.51. That brought the total cost, before taxes, to $14.31, or nearly double the price Amazon had listed on the initial page.

**What kind of sophisticated shopping algorithm steers customers to a product that costs so much more than seemingly comparable alternatives?**

***Hypothesis/assumptions:***

* Amazon manipulate prices and ranking products .
* products appear in buy box are owned for Amazon or vendors that have subscribed to ( FBA ) programme.
* The ranking of products was biased on the price it doesn’t list shipping charges for products sold by Amazon or FBA during browsing but adds in the last step.
* Other vendors that are not subscribe for FBA there chance of appearing is less or poor, even if they meets the requirements.

**Data and Design**

This dataset includes:

* **ScrapeDate**
* **ProductName**
* **Ranking**: ( ScrapedIndex, ScrapedIndexTrueRank, CorrectedIndexTrueRank )
* **Price** ( ScrapedIndexPrice, CorrectedPrice, BBCorrectedPrice )
* **Vendor**( ScrapedIndexVendor, BBVendorScrapedIndexVendorType, BBVendor )

***Models***

Logistic regression

**Tools :**

* Numpy and Pandas
* Scikit-learn
* Matplotlib and Seaborn