

**Data Analysis of Online sales around the world**

**Submitted by**

Mr. Sreerama Maram

**Registration No**: 11610529

**Roll no:** A25

**Section**: EEM34

**Course Code:** INT 217 (Introduction to Data Management)

**Under the Guidance of**

Mr. Parminder Singh

**Discipline of ECE**

**Lovely School of Electrical and Electronics Engineering**

**Lovely Professional University, Phagwara.**

**DECLARATION**

I, Sreerama Maram, student of Lovely Professional University (Program name) under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 11 Nov 2018 Signature

Registration No. 11610529 Name of the student: Sreerama Maram

**CERTIFICATE**

This is to certify that Sreerama Maram (bearing Registration no. 11610529 has completed INT 217 project titled, “Data Analysis of Online sales around the world**”**

under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

**Signature and Name of the Supervisor**

**Designation of the Supervisor**

**School of Electronis and Electrical Engineering**

Lovely Professional University

Phagwara, Punjab.

Date:

**Introduction:**

Online sales is one of those big business categories which have gone through a major turnaround during the last decade. It has impacted, either directly or indirectly, various departments of business and even non-business models. Some have benefited a lot through it and many a few, unfortunately, suffered.

Nonetheless, right now we (I literally mean the whole world) are moving into the state of the business model where even almost all the physical products manufactured are to be sold online. With everyday groceries to Tesla’s racing cars being delivered to our home faster than we can get them personally, the end of offline seller business is clearly inevitable. With something like this around the corner, which is potentially an end for offline stores and sellers, we are in need to realize the best tactics to follow for a smooth, less error prone, and most importantly an efficient process of selling any kind of goods online.

So, let’s try and do it.

**Scope of analysis:**

The data used in this analysis is a detailed record of more than 500,000 orders which were placed round the world during the last decade (precisely 2010-17). Using this we will be trying to find any patterns, rises, down falls, and possible relations between different aspects and try to conclude a real-life usage of the findings.

**Existing System:**

Nothing seems wrong until something really better comes along.

If nobody started something like carpool share rides, obviously there wouldn’t be any complaints, but only when we start enjoying the new, we start dwelling about the part.

Who would have thought that Gotham could use the batman? But when he did rise, everything became so cool and simple. The purpose of this research is not to really start with a problem and try to solve it, but to find any new insights that the online business sellers were missing on all these days and use them for a better progress in sales.

The aspects we will be concentrating about are:

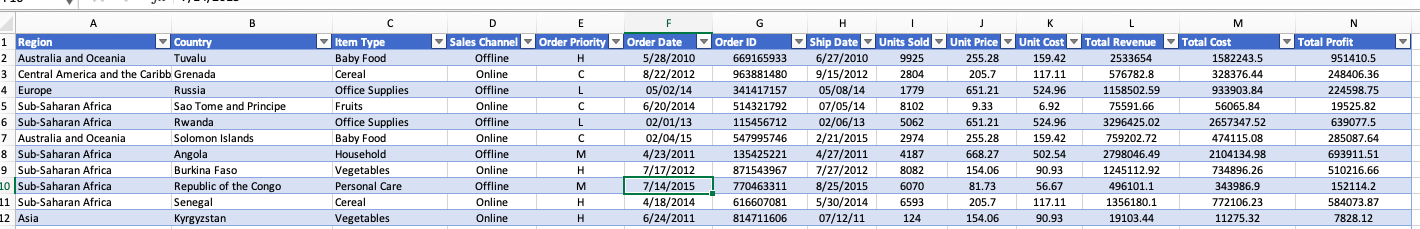
* How the no. of online sales increased or decreased during the last decade in different continents and countries.
* The increase/decrease in net profit of different categories of products (useful to know which category of products sell best online)
* The demand for order priority.
* The increase/decrease in delivery speed as the years passed.
* Review whether having order priority greater(thereby paying more money) really benefited the customer.

**Source of Data:**

Thanks to [eforexcel.com](http://eforexcel.com)

link: <http://eforexcel.com/wp/downloads-18-sample-csv-files-data-sets-for-testing-sales/>

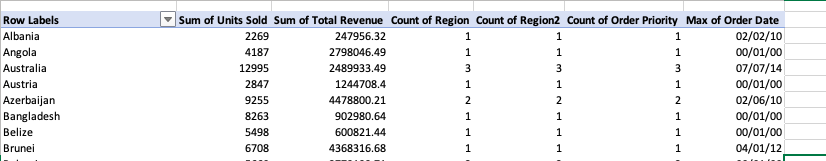
Sample view of Dataset:

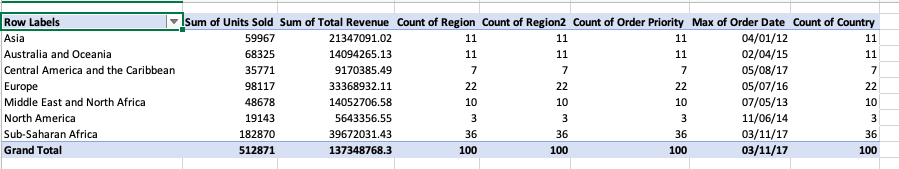


**Data Analysis:**

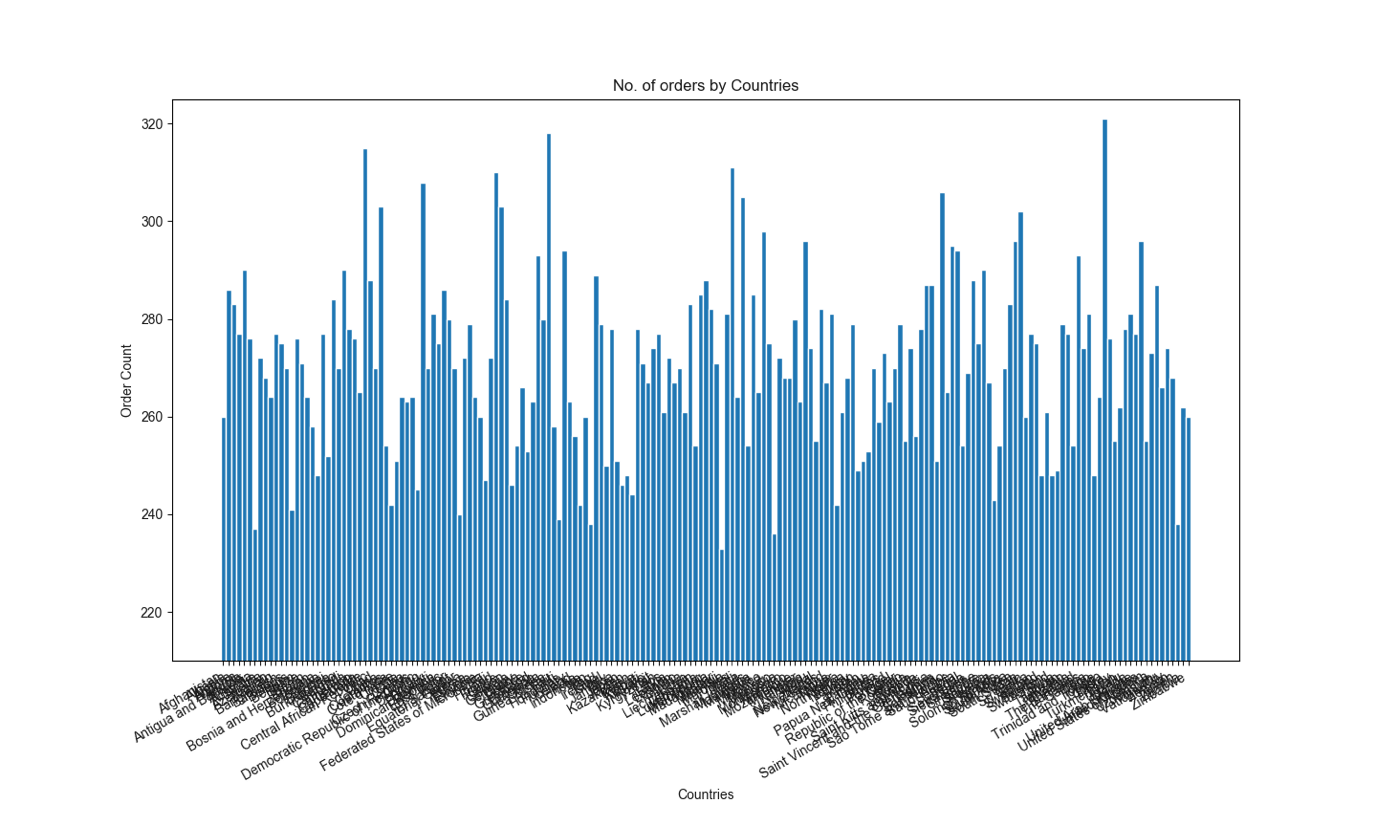
**Rise/Fall in online sales:**

Let us begin with grouping by the order data by countries and continents to get find the trend in online business region wise. The pivot table below shows the same:



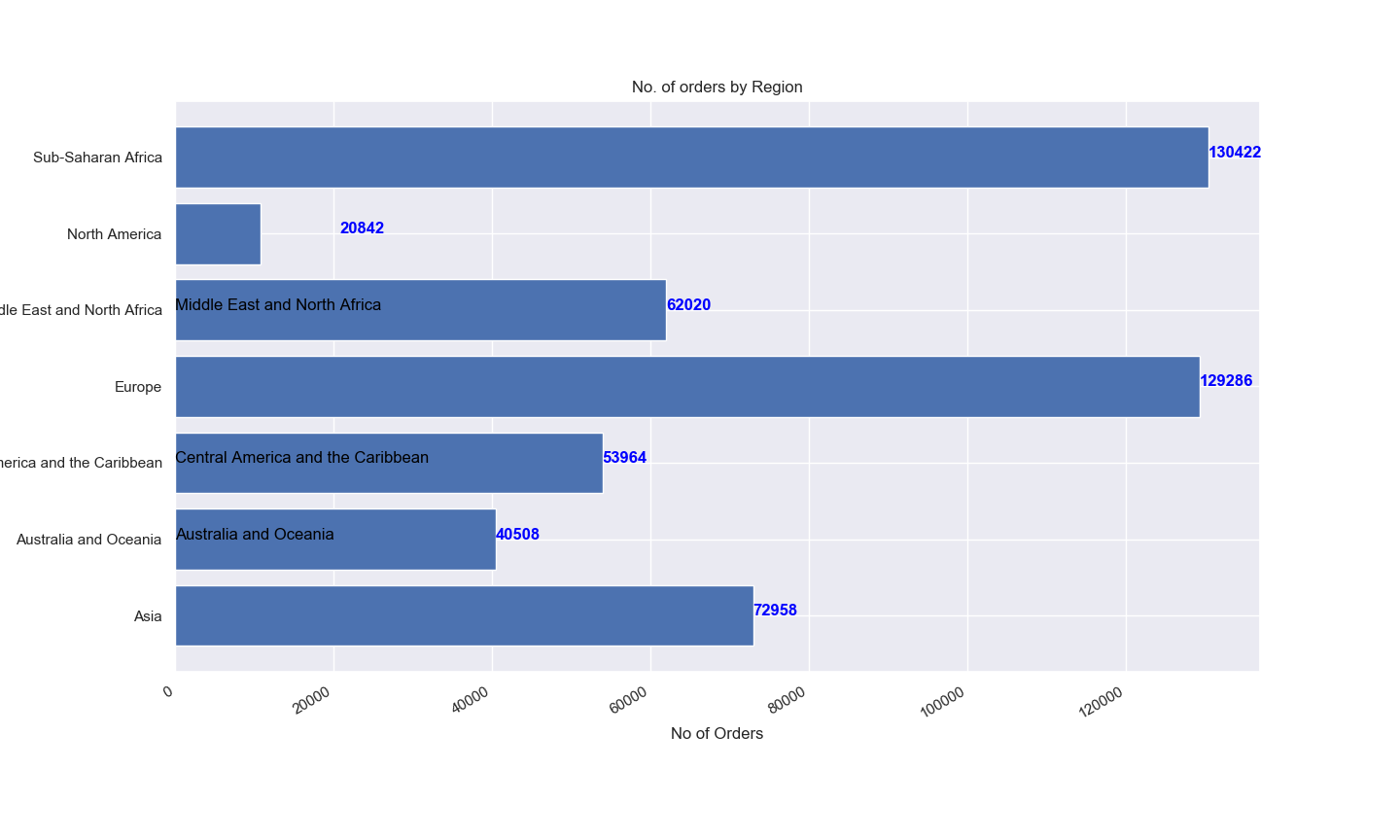


The country and region wise revenue can be seen clearly of the sample data. The graph between then is shown as follows:



A zoomed look into this graph would clearly show that the order count increases lot depending on the population of a country, and it’s technology wise development. This is concluded since countries such as Kazakhstan and the ones in Central Africa are seen to less order count while the United states of America the one of the highest order counts.

Having a look at a graph on Region wise order count:

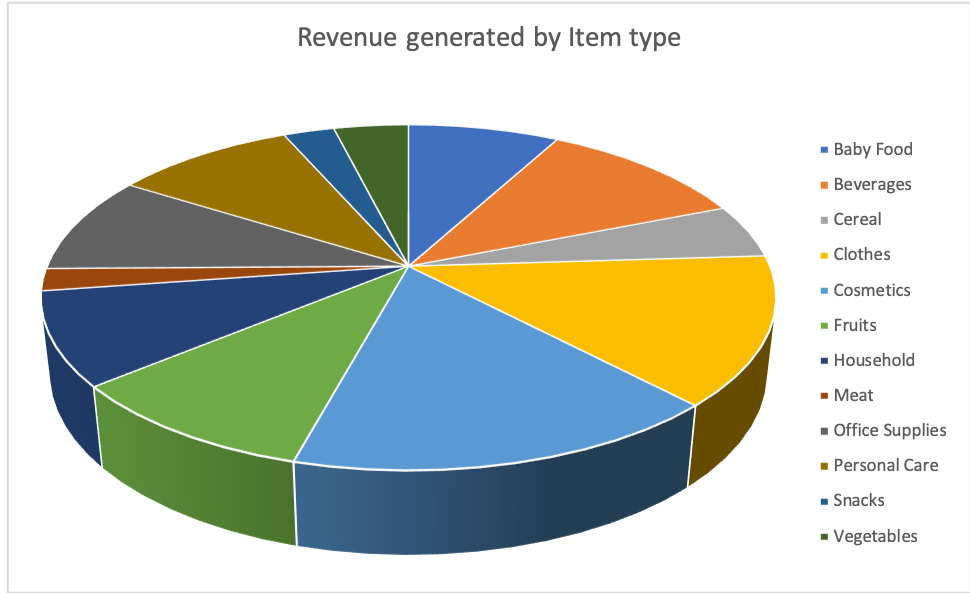


It turns out that the number of countries in a continent really affects the number of orders placed, Asia being the biggest continent has less order count than Europe which has a number of smaller countries and which is also better development wise.

It conclusive that aiming an online business in developed and divided regions like Europe and Sahara Africa is better than doing it in regions like Asia. The reason for this could be that since regions as Europe are so finely divided it is not easy to buy anything one wants from the same countries. But in case of countries such as India which is geologically a sub-continent and bigger countries as the United states of America, online sales did get effected since buying a product there is not as big a problem.

**Trend of sales by Item type:**

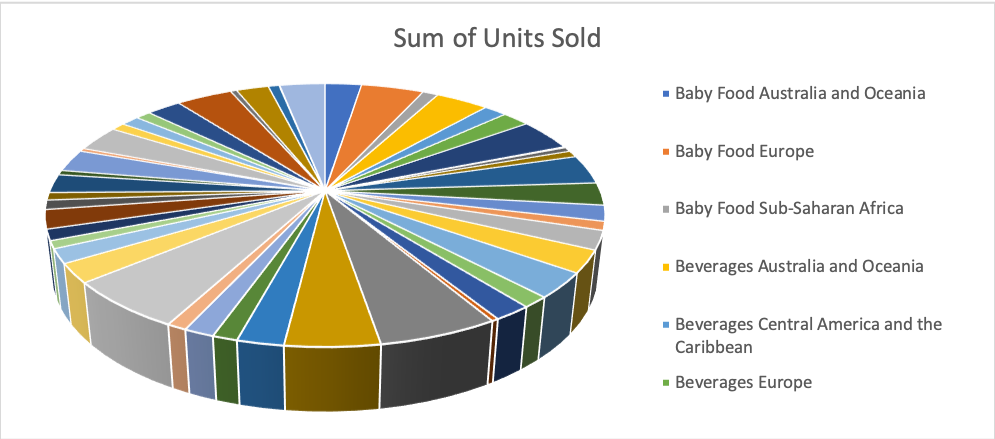
Deciding on what items to sell and to know on what item to put an extra effort is very important while doing an online sales business. Let us take our data and see what kind of products sold themselves the most during the last decade.



Very surprisingly, Baby food has taken the number one spot in revenue generation. Nobody would have expected this, but this could be because of the high price of baby food and supplies and since people give quality of baby products a big priority thereby not really caring about the price of the products.

So starting an online baby products only sales site could be a great money generating idea.

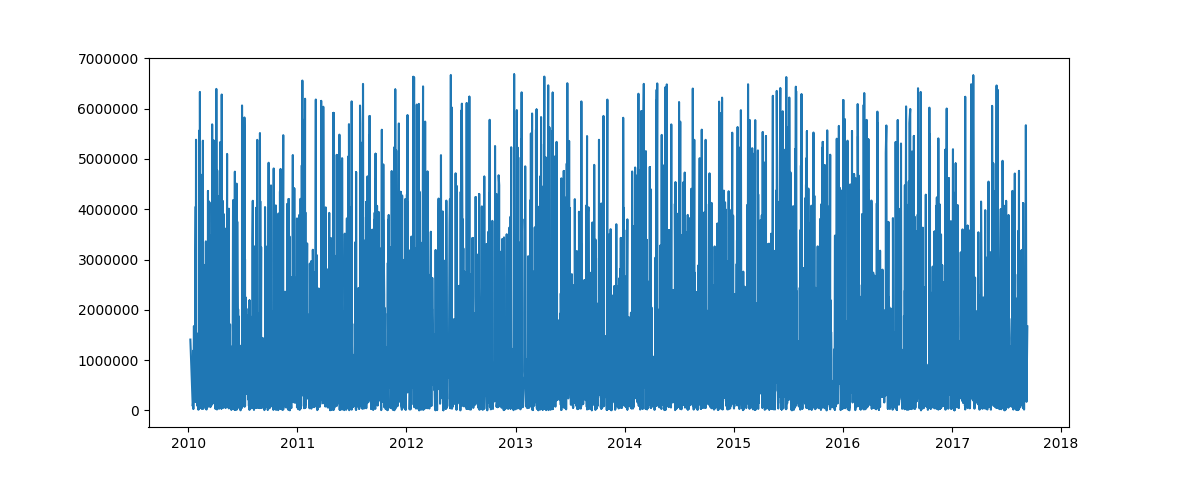
And not so surprisingly, vegetables and clothing takes the next two spots, because food is must for human, and shopping is must for woman. Meat takes the last spot because of those vegetarian people( that doesn’t mean not to start a meat shop, do it, I like meat).



**Rise in revenue by date:**

It is no doubt that the online business has increased lot during the last decade, but in spite of this we are not seeing a big rise in revenue of sellers. This is surprising, but on further research I found that,

The online business has rose much in the underdeveloped countries if Asia and African in the last few years(shown in the first graph). So since the data of the revenue we have in our data set is calculated in US dollars, it sure must had an effect on the revenue graph. So it can be concluded that even thus plain straight plot is a good sign of rise in number of online sales around the world.

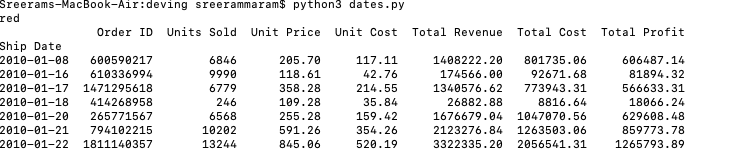


**Effect of Order Priority:**

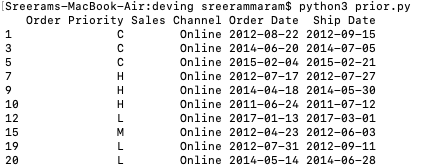
How many times did it happen that you order something online and it gets delivered to you the very next day or inside two days, even without paying for the higher priority order. Almost all the sellers these days deliver the good at a very fast pace to the buyers, but we still see the costumers who pay extra for order priority just to make sure something delivers fast.

In this module, lets test whether the consumers paying for higher order priority is worth it or not. Is it really having an effect on the ‘delivered date’ or not.

As we group our data set by date:



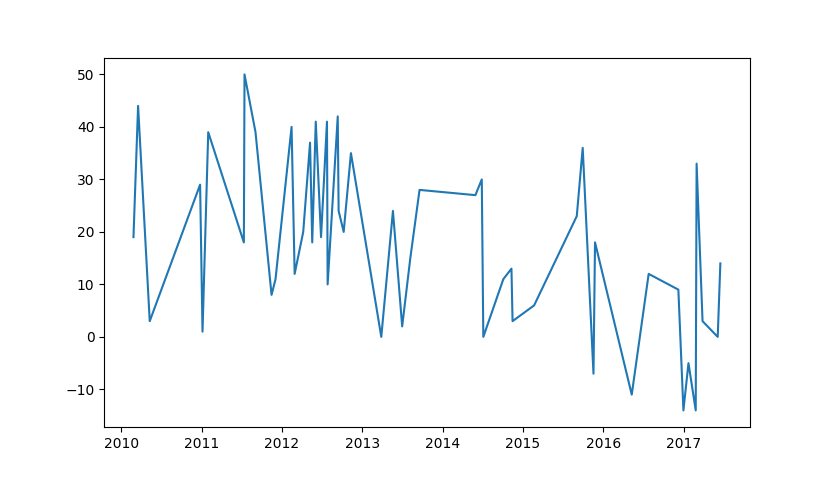
Just care about the order dates and priority:



And add a column which shows the difference in number of days between the order date and shipment date:

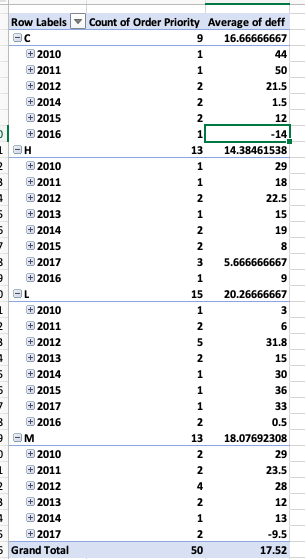


Now lets plot a graph showing the change in delivery period:



Which clearly shows a big decrase in delivery time a the years progress.

The below pivot table shows the average time take in days for the product to be delivered over the years for different priority types:



A big difference is found in pace of delivery as the years passed. More importantly it can be send that having a medium priority rather than a low priority gets the product delivered a lot faster but there is not a significant difference when H and M priorities are compared.

**Future Scope:**

While a big boom in online business is on it’s way, it becomes important for both the sellers and the buyers to find new and efficient way of selling the product and also buying smartly.

Understanding which category of product are going to sell the most, and generate more revenue( like we found that baby products generate huge profits) would help the sellers better their business. It is also important to aim the right regions and societies for the making the sales as high as possible. Analysing the past data to make the right decisions will definitely help the business model.

It is also important, as the business progresses, to look at the features of it which are still impactful and the ones that are need to be dropped which doesn’t effect in improving the business but only wastes the money of the buyer( like in case of priority orders)

**References:**

<https://www.bdc.ca/en/articles-tools/marketing-sales-export/sales/pages/3-keys-online-sales-success.aspx>

<https://www.sciencedirect.com/science/article/pii/S0167718717300929>

<http://eforexcel.com/wp/downloads-18-sample-csv-files-data-sets-for-testing-sales/>

<https://stackoverflow.com/>