Candidate number: **190533784**

**Analytical report**

**Table of content:**

**Executive summary**

Data that is under investigation is provided from an international internet shop that requires analytical analysis. The dataset itself consists of 24 columns and approximately 51 thousands of rows. Due to the wide range of the parameters recorded, analysis of the data as a whole will not be consistent and of full value. Therefore, in order to provide full-filled analysis, one would be more likely to succeed by reducing the number of columns under investigation. One of such approximations for the analysis is described below on the topic of discount allocation for the sales department manager.

Reliable shop business requires stable sales which are usually stimulated by discounts for products. In most cases discounts are used to stimulate the so-called willingness to pay metric which is not under consideration in this report. However, the discount strategy of that particular shop is not obvious from the first glance. The goal of the report is to identify the approximation of the pattern by which the discounts are allocated in the current state and suggest improvements on that strategy, which should optimize profit.

There are two possible strategies on the path to profit optimization. One of such is increasing loyalty of customers, in order to increase the probability of their return to the shop.. What is more important is that those customers have to be with profitable orders, since purchases with negative profit are also in our range of profits. Another possible way is to decrease costs, in terms of costs that appear with discount establishments. First rule here is suggested on the basis of shipping costs. Reduce discounts for products with high shipping cost. Second rule is to allocate bigger discounts to the orders with the lowest time requirements.

The whole analysis, depicted by 5 dashboards, is logically divided into the spheres of the discount applicability. Those are: Geographic dashboard, Client's dashboard, Product's dashboard, Order's dashboard. The last dashboard shows ambiguous relation of the discounts that was found in the data. It is called Order's relations. Analysis is done on the whole data set and due to its size time periods analysis was not conducted. All the dashboards are provided below their description as well as in the separate document. Each figure has its description, since tableau doesn't provide needed legend and that information would be included in the oral presentation of the report.

Due to the lack of information about the dataset, few assumptions were made. More precisely, discounts are listed in the form of sales share. Then later, a new calculated metric of absolute discount is used, which is simply the sales multiplied by discounts for each row. That was made in order to aggregate different groups of objects with the sum method applied to discount without losing inference ability. Further in the report by discount always meant absolute discount. Also, shipping costs are assumed to be costs of the shop.

**Geographic dashboard**

Starting from the geographical aspects of discount strategies at the top of the dashboard, one is able to depict relation between sales and discounts on the basis of location change.

Figure description:

It is a map of all countries of the world, on which countries, in which delivery is conducted, are labeled with colours. The colour explains the sum of absolute discount that is allocated to a certain country. Every country has a dot at the center, the size of which is related to the sum of sales in it.

Meaning description:

Fast label of top sale countries: US, Australia, India, China etc.

Fast label of top discount countries: US, Australia, France, Mexico.

Top is similar, but the diagram shows inconsistency. Proper picture here would be bigger dots in the countries that are coloured in dark green, where discounts are proportional to sales. However there are such countries as China, India, Brazil and so on, where dots are relatively big and their colour is more yellow than green. That fact says, that discounts and sales in countries are not proportional

In the interest of that inconsistency another chart was constructed on the left bottom corner of the dashboard.

Figure description

Bar chart displays the new calculated metric "Discount sales ratio" (DS ratio further), which is indeed the absolute discount value summed for each country and divided by the sum of sales of that country. It explains how many dollars of sales are there for one dollar of discounts made for products and orders in that country. There are top 5 countries from top and tail of sorted countries list by DS ratio. Average sales and discounts are presented in bar charts.

Meaning description

Such a metric was established in order to find countries where sales are high without discounts and on the other hand those where discounts are high, but sales are not.

A different passage is in the chart on the right bottom of the dashboard, which tries to visualize shipping costs.

Figure description

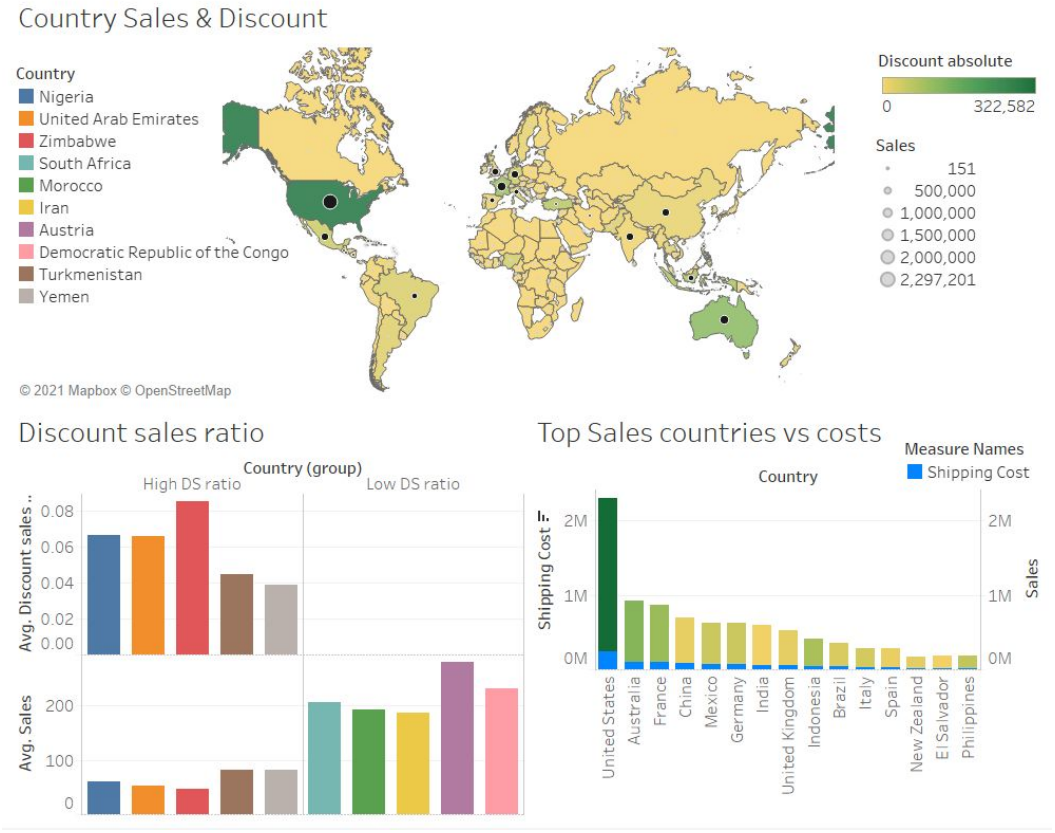
15 countries with the highest sales have their shipping costs and sales on the y axis. Colour of each bar explains the discounts sum.

Meaning description

Countries with high sales sum have proportional shipping costs to sales. However discounts are more or less random.

Dashboard conclusion

Geolocation aggregation doesn't show relation between sales sum and discounts sum. Moreover there are countries with low discounts that have higher sales on average.




**Client's dashboard**

Client analysis is crucial in any business. Second Dashboard consists of one diagram which is called Loyalty vs Profit.

<u>Figure description</u>

It is a coordinate plane with profit on the x axis and loyalty on the y axis. Loyalty is the number of times, which a client is mentioned in the data. Those might be mentions related to different or one order with many positions in it. However after 45 mentions in the data, a client becomes loyal. Labels on the plane describe clients. In each quadrant clients have different labels. Colour of the label shows the sum of discount that was allocated to that client.

<u>Meaning description</u>

Simple separation of the clients is allowed by the diagram. It is possible to depict loyal and profitable clients as well as unprofitable and unloyal. That separation should be one of the crucial rules for discount allocation. Moreover, for current improvement it is possible to find clients, which have high discount levels and are in the third quadrant of unloyal and unprofitable clients.

Dashboard conclusion

Clients analysis provides evidence for a positive relation between loyalty and discount, but an absence of relation between discount and profit.


**Product's dashboard**

Since the business is working with multiple categories of products, it might be beneficial to start analysis of product discounts with a pie diagram at the top.

Figure description

It displays the variety of subcategories as a colour within the share of absolute discount sum as an angle of a pie chart. Outer shell provides names of categories.

Meaning description

One may see all the categories of the shop and share of discount for any subcategory. Listed subcategories have the biggest discount.

In the left bottom corner Products are seen from the shipping cost point of view.

Figure description

It is a bubble chart where each bubble is a different subcategory. Size of bubbles is bigger with a higher average shipping cost. Colour is darker with a higher average discount.

Meaning description

Here the relation between the shipping cost and discount is seen. The higher the average discount the greater the average cost. On the basis of assumption about shipping costs, in order to decrease costs it is wiser to stimulate products with low shipping costs. Current situation is 180 degrees to the wrong way.

Last diagram on the dashboard works with the quantity of sales. (Bottom right)

Figure description

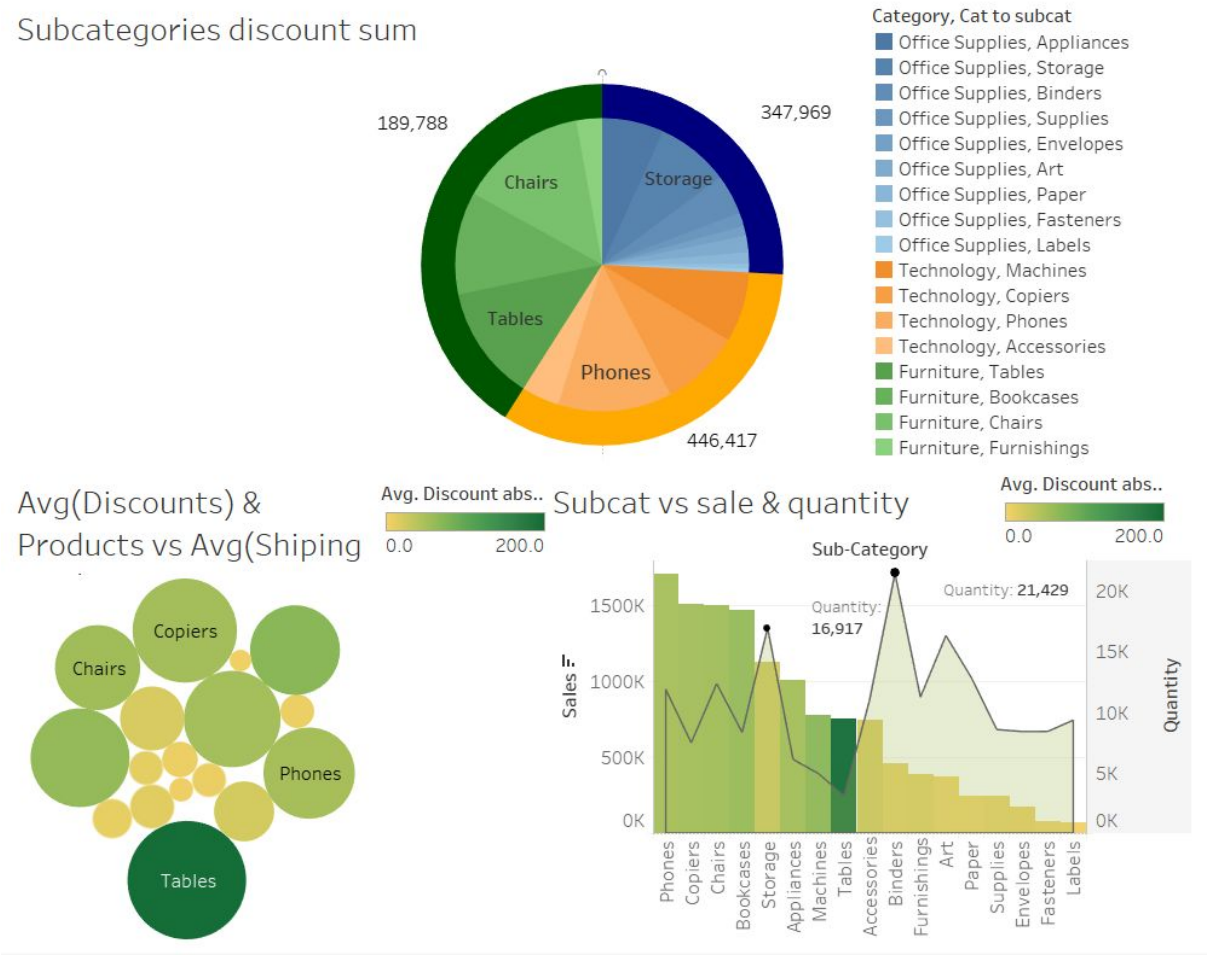
Bar chart of subcategories sales with area graph of sold products quantity on the same axes. Colour of the bar chart is related to the average absolute discount. Subcategories are sorted in sales order.

Meaning description

Left half of the diagram with higher sold products has generally higher discounts and lower quantity of sold products. Such a relation is understandable. Products that bring more money are promoted by discounts and products which are sold in high quantities, but don’t bring sales are not stimulated by discounts.

Dashboard conclusion

Product aggregation helps to depict positive relation between sales sum and average discounts, negative relation between discounts and quantity of products sold, which are understandable. Positive relation, between average shipping costs and discounts, which is unneeded, but exists.




**Order's Dashboard**

Orders are the only entity that has additional time relation, which is called ship modes. Therefore let's start analysis of discounts in orders with the diagram on the left.

Figure description

Sequence of histograms, which are stacked vertically in accordance to the fastest ship mode at the top. Each histogram shows the number of orders divided in 5 groups different by profit value. Colour says the average value of discounts for those orders.

Meaning description

One may observe that orders around 0 profit are the most common as well as orders of standard class. In each class discounts are distributed in the same way, however it is advised to leave discounts only for orders in the standard mode in order to decrease expenses related to the short timing. In addition, unprofitable orders have an inadequate amount of discounts which also have to be fixed.

On the right top corner there is a simple figure of orders sales and discounts relation.

Figure description

A treemap diagram of 100 orders with highest sales. Size of the boxes describes the sale variable, colour is responsible for the discount sum.

Meaning description

Discount sum and sales sum of the orders are not related among the orders with highest sales value. Shipping cost of the products are summed up into the shipping cost of order.

Last diagram at the right bottom corner starts the investigation on the relation between shipping cost and discount amount.

Figure description

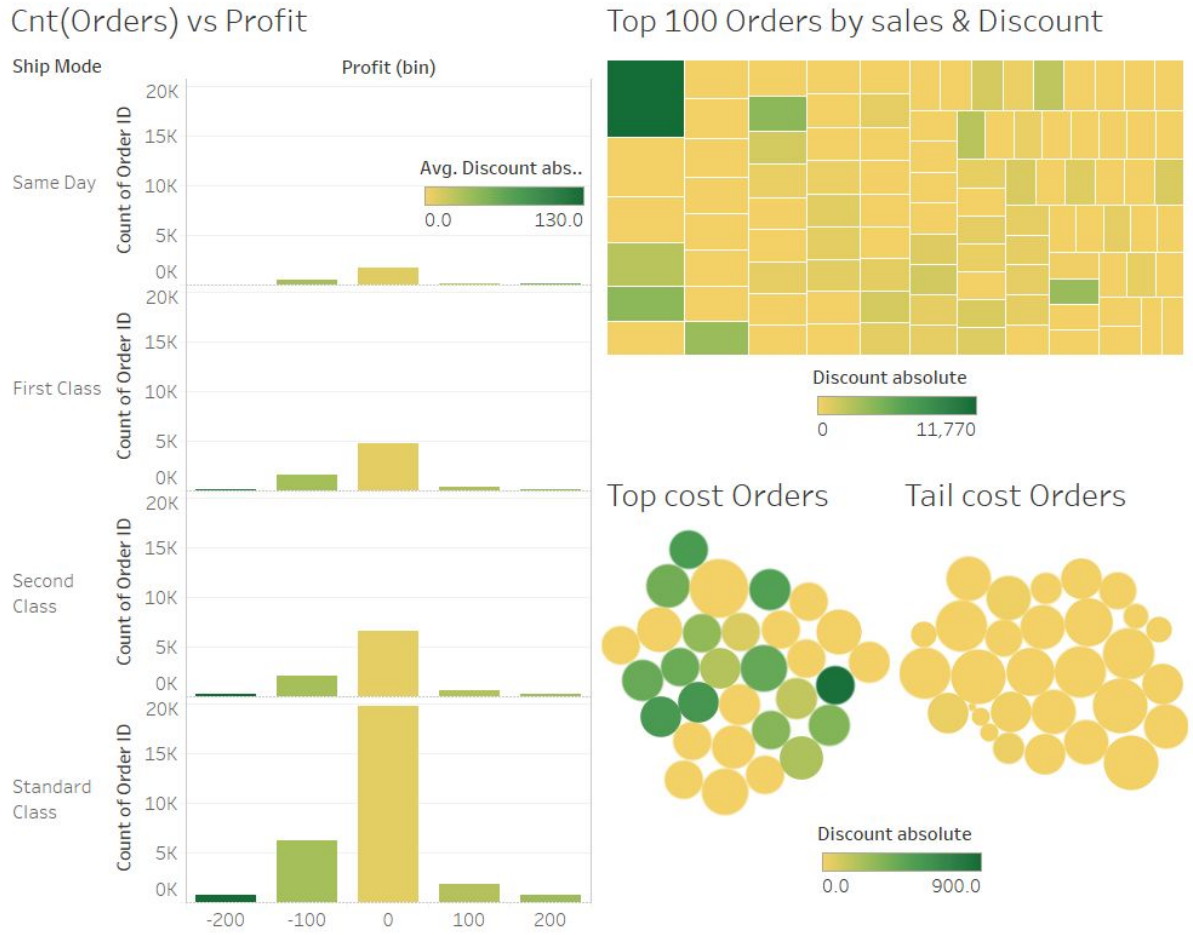
Comparison of two packed bubble diagrams placed side by side where each bubble is an order. Left one describes the top 30 of orders with the greatest shipping cost, while the right one describes the top 30 of the lowest shipping cost. Their sizes depict shipping cost without dual axes, since bubbles on the right side would not be seen. Colour shows the discount value for the order.

Meaning description

There is some positive relation since on the left side of the figure green circles are common and on the right there are none of them.

Dashboard conclusion

Analysis of orders provides information about the lack of relation between ship mode and discount which is advised to be added. It is also suggested to drop discounts for unprofitable orders which is common and make further investigation on the relation between shipping cost and discount.




**Order's relations**

There is a negative relation of profit and discounts of orders that was mentioned on the previous dashboard, as well as a positive relation with profit. Both of them are under investigation in the following two figures.

Figure description

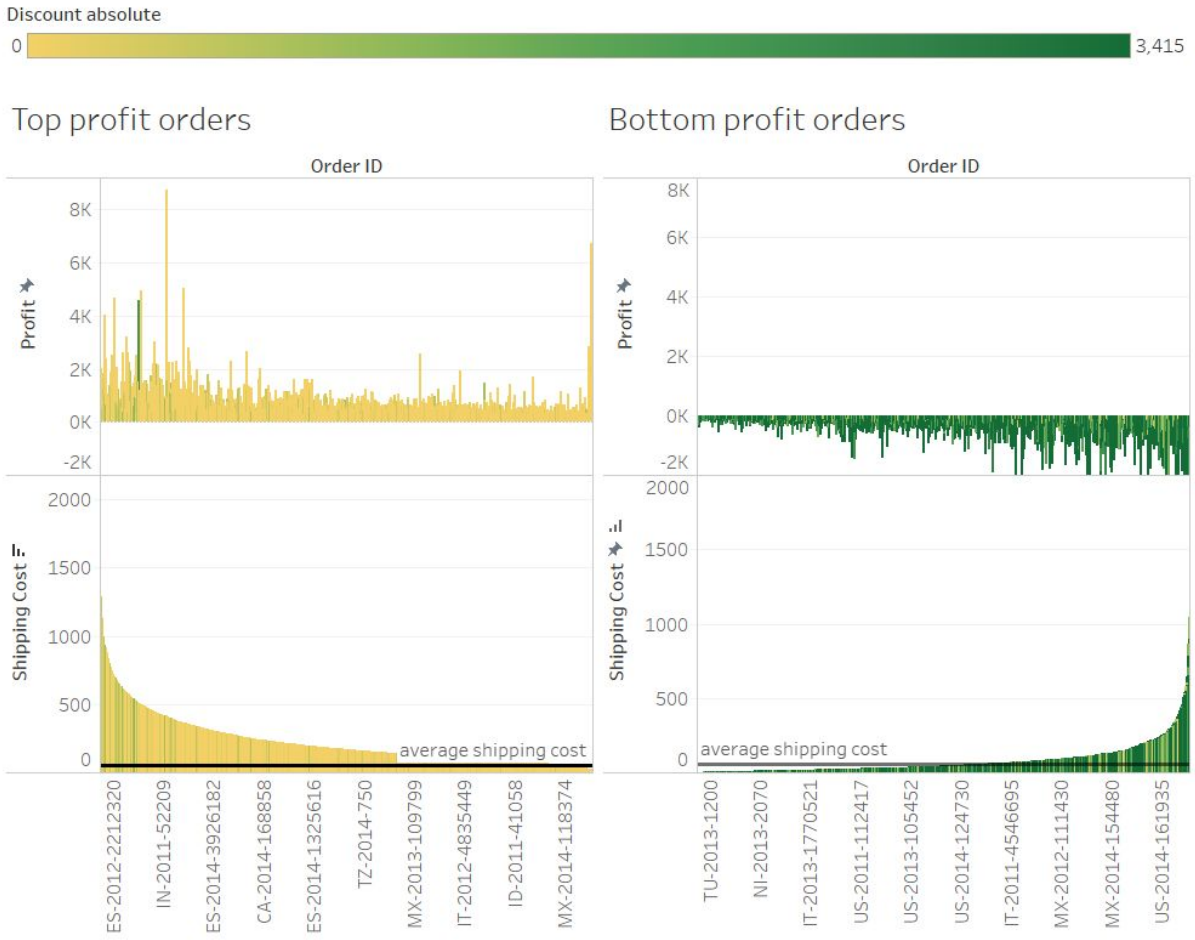
Two bar charts of 2000 orders with highest profit and lowest on the x axis on the left and right respectively. Both couples depict Profit and shipping cost on the y axis and discount as a colour parameter. Everything is sorted by shipping cost.

Meaning description

From the charts can be seen the negative relation between discount and profit. What is of bigger interest is the relation of profit and shipping costs through discount. Orders with high by module profit have high shipping cost and polar discounts.

Dashboard conclusion

High variability in profit has a positive relation with shipping cost, but no relation with discount. After implementing all the advice listed in the report the relation should be found and provided as a result of the discount allocation report.




Word count: 1488