Revolutionizing Fashion: Leveraging Data-Driven Solutions for Anvi Be Yourself

A Mid-Term submission report for the BDM capstone Project

Submitted by

Harsh Anish Shah

21f3001359

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

TABLE OF CONTENTS

Revolutionizing Fashion: Leveraging Data-Driven Solutions for Anvi Be Yourself

1 EXECUTIVE SUMMARY

ANVI -Be Yourself is a fashion brand that has its presence on various e-commerce platforms. The company aims at selling quality fashion products all over the country and plans to expand its presence in the coming years.

There are some business issues that the company is facing and one of them is related to inventory management. As the company is related to the fashion industry, the trends keep changing on a regular basis and hence the company is also facing issues with what the “age of the product” should be. The company also faces issues related to Product performance and hence product based performance analysis also needs to be looked into.

The issues will be addressed by analyzing the data via some of the analytical approaches to obtain an outcome. For the issue related to inventory management, demand forecasting is something to be looked at by analyzing sales data and external factors to predict demand accurately and hence help optimize inventory levels. And for the “age of the product” issue, I am looking at Trend analysis by looking at fashion blogs, customer reviews to get to know emerging trends and then collect them and try to include the same in new trends.

The expected outcome will help solve issues that the company is facing and will help optimize their business and increase profitability of the company.

2 PROOF OF ORIGINALITY OF DATA

LINK-
https://docs.google.com/document/d/1mTjggAr4YsteRyk_XWr_xe_FLR1aKpid/edit?usp=sharing&ouid=102986606727583424316&rtpof=true&sd=true

Also attached towards the end

3 METADATA

Metadata comes in various types, each serving a specific purpose to provide context, organization, and understanding the data. Here are some types of metadata that I have come across while using the company’s data for analysis:

3.1 PRODUCT METADATA:

- Products: The company currently has around 500 different fashion products which have been rolled out in the market and all of these are present in the data under SKU names.
- Fabric : Fashion industry is very much classified according to the fabric that the clothes are made off and hence they have an entire dataset on different fabrics that are being used. From fabric quality to fabric shade, from type of fabric to the amount of the fabric ordered for manufacturing.
- Colors and Variation: Under the fabric dataset itself comes the different colors and variation that they have been using for making a particular product. They have around 250 different shades and color variations for their products.

3.2 SALES METADATA:

- Sales Data: The company has a wide collection of data of all the customers along with the product purchased, size, region from where the customer is based at, sale price along with the discounts given and the e-commerce platform as well on which the purchase was made.
- Customer Feedback: Attached with each product bought is a customer review, and for the current time being they have only logged in negative reviews.

3.3 INVENTORY AND LOGISTICS METADATA

- Inventory Status: This data contains total stock in hand, amount of stock sold, amount of stock required along with the SKU code.
- Logistics: This data contains the information about the courier agents with whom the delivery of a product is scheduled
- Stock Data: This data contains parent SKU, SKU, Manufacturer name, QC check, Fabric costing, Fabric consumption data.

4 DESCRIPTIVE STATISTICS

Descriptive Statistics can provide valuable insights into a fashion company's dataset and it will help to summarize and understand the data.

4.1 SALES DATASET

SELLING PRICE	VALUE
MEAN VALUE OF SELLING PRICE	819
MEDAIN VALUE OF SELLING PRICE	795
MAX VALUE OF SELLING PRICE	3300
MIN VALUE OF SELLING PRICE	383
RANGE OF SELIING PRICE	2917
VARIANCE IN SELLING PRICE	44261.65
STANDARD DEVIATION OF SELLING PRICE	210.3845

Figure 2 : Selling Price Statistics

SKU NAME	
MODE ITEM OF SKU AVILABLE(Most Selling)	Brown - Black Top
2nd Most Selling SKU	ANVI Be Yourself Women Boxy OpaqueTie Ups Casual Shirt

Figure 3 : SKU Statistics

SIZE	
DIFFERENT SIZES AVAILABLE	XS,S,M,L,XL,XXL
COUNT OF S	12449
COUNT OF M	15563
COUNT OF L	11951
COUNT OF XL	7932
COUNT OF XXL	6874

Figure 4 : Different Sizes Statistics

CHANNEL NAME	
DIFFERENT CHANNELS AVAILABLE	AMAZON_IN_API, an
AMAZON	725
FLIPKART	4351
BEYOND FOLLOW	12
MYNTRAAPPMP	28090
NYKAA	14707

Figure 5 : E-commerce Statistics

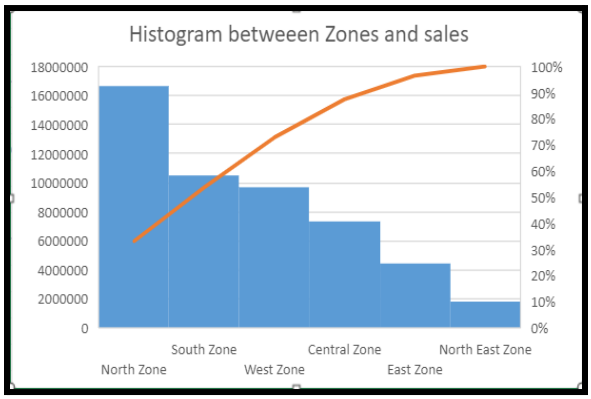


Figure 6 : Histogram of Zones vs Sales

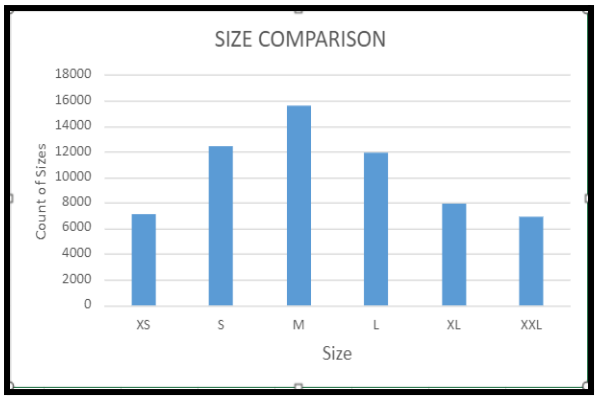


Figure 7 : Bar Graph of Size vs Total Stock

- The sales data contains most of the data that will be required for the analysis of the problems that were mentioned in the problem statement earlier.
- After cleaning the data and having a good look at the data, I think that the **selling price** parameter is a very important parameter for further analysis of the solution to the problem statement.
- SKU Names(The product names) have been shown above as the best and the second best ones but it has been segregated in descending order as well in order to know more about the product performance.
- Product further leads me to the point that the size of the products have been more in demand and hence performing some statistical analysis for the “size” parameter may also be helpful. A Line graph is plotted as well: denoting which size is more in demand in order to derive basic conclusions.
- E-commerce platform sales is another parameter that I have got some statistics for and this will be quite beneficial for promotion of the products more on the best performing firms.

4.2 FABRIC DATASET

FABRIC		
FABRIC AVAILABLE	81380.07	meters
AVERAGE FABRIC AGEING	406	
TOTAL AMOUNT SPENT ON FABRIC	#N/A	

Figure 6 : Fabric Statistics

- The fabric dataset is the most primary data that has to be kept track of and after which the outcome of the number of clothes can be considered . The total fabric available is quite high and this data is only for the first half of the year 2023.
- Fabric Aging denotes how long the fabric is kept in the inventory for making the product and the aging is not quite high and is neither quite low but it does indicate that it is more than a year which could be reduced depending on the inventory reduction by more sales.

4.3 INVENTORY DATASET

INVENTORY	
TOTAL STOCK	56224
AVAILABLE IN HAND	56177
MAX STOCK FOR A PARTICULAR ITEM	645
MIN STOCK FOR A PARTICULAR ITEM	1
AVERAGE STOCK KEPT FOR AN ITEM	29

Figure 7 : Inventory Statistics

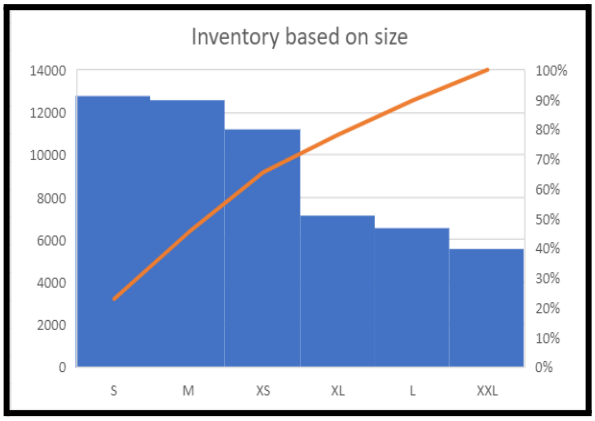


Figure 8 : Histogram of inventory based on size

- The inventory dataset is very important as inventory management is also one of the parameters that I am looking to improve and hence basic statistics has been shown above for the same and will be further creating Pivot tables based on the inventory data along with different filters in order to get appropriate results.
- A histogram is also plotted in order to get a basic idea about how the inventory is distributed over various sizes.

5 ANALYSIS PROCESS/METHOD

The analysis process and methods for this project differs according to the problem statement and the data available related to that particular problem.

Looking at the customer segmentation issue, the dataset has been divided based on the 6 zones according to the Indian map where each zone contains at least 4 states.

- North Zone, East Zone, North-East Zone, West Zone, South Zone, Central Zone

- Further, Months have been divided in the form of quarters and then segregated based on the quantity sold in these regions. According to what I feel, if both the quarters show an increase or almost equal demand based on the region, it can be conclusive that more customers are attracted to the product from that particular region and hence more shadow has to be put on the customers from these zones.
- Of course, there might be external factors like population as well to be taken into consideration but still according to the analysis, even if North east and the East zone is combined, it is almost equal to the segmentation of the central zone(quantity sold). Hence, this is one of the approaches that is being followed.
- Furthermore, looking at which product is the most sold in the highest selling region would further aid in looking at the profitability issue by trying to get more of the fabric used to make the high selling products and hence reducing inventory and fabric aging as well.

Count of QTY	QUARTER		
ZONE	Q1	Q2	Grand Total
Central Zone	4517	4377	8894
East Zone	2770	2573	5343
North East Zone	1105	1079	2184
North Zone	10287	10884	21171
South Zone	6391	6119	12510
West Zone	5788	5810	11598
Grand Total	30858	30842	61700

Figure 9 : Zone wise segmentation

Size	Count of SKU Name
L	11924
M	15518
S	12402
XL	7914
XS	7055
XXL	6857
XXS	6
Grand Total	61676

Figure 10 : Size wise segmentation

- One more factor that is being considered for customer segmentation is based on the most popular size and hence it suggests a solution of more production of the size that is more in demand.
- In my opinion, Customer segmentation leads to increasing profitability because of the information of knowing what type of products and where we would be trying to target to try improving sales.
- Improving profitability and sustaining is one of the most important considerations of any successful business and hence in this case as the company is an e-commerce based fashion business, segregation of the sales of all the products across various platforms is performed.

- There are more than 5-6 platforms taken into consideration and hence it becomes a solid consideration for better marketing and hence more profitability. Interesting fact to be noted here is that by trying to focus on the platforms where more sales have been done, consumer behavior can be noticed as well and hence it contributes as a factor for the Customer Segmentation issue as well.

Channel Name	Values									
	AJIO		AMAZON_IN_API		Anvibys		BEYOND FOLLOW		FLIPKART	
Month	Sum of Selling Price	Sum of QTY	Sum of Selling Price	Sum of QTY	Sum of Selling Price	Sum of QTY	Sum of Selling Price	Sum of QTY	Sum of Selling Price	Sum of QTY
Jan	784338.62	1031	82049	73			6243	7	469544	604
Feb	716062.67	991	68472.14	74	3465.35	3	2100	2	592866	804
Mar	1671025.37	2384	195113.38	237	5561	5			941423	1405
Apr	1180527.99	1698	121332.34	141					501433	670
May	1189319.36	1640	103159.16	119			490	1	415870	513
Jun	853811.17	1198	66462.16	81	1590	1	1433.6	2	315203	355
Grand Total	6395085.18	8942	636588.18	725	10616.35	9	10266.6	12	3236339	4351

Figure 11 : Platform wise sales and quantity sold

- The above chart also takes into consideration month wise sales and quantity sold and hence this analysis could be further broken down into why a particular month was performing better across all the platforms.(eg. Holi(A hindu festival) brings in a huge sale on these platforms and hence for the few platforms shown above out of the many, notice can be taken of great sales in March.)
- Keeping a great inventory in these months is very critical and hence this leads to the next issue of taking stock of the Inventory available. A very comprehensive analysis technique is trying to be used; firstly: by segregating each product according to its size and finding out how much stock is available. Then moving on to the sales data and seeing the previous year or quarter wise data of the previous year and finding out if the inventory currently present is at least equal or more than the previous year's stock present especially in the high selling zones that had been segregated earlier.

Sum of Total Inventory	SIZE-letter												
Item Name	1	3	4	5	L	M	S	XL	XS	XXL	XXS	(blank)	Grand Total
Red Floral Printed Top					408	645	532	209	204	206			2204
Brown - Black Top					361	417	407	152	214	139			1690
Mint Green Floral Printed Top					297	408	341	168	273	141			1628
Stylish Yellow Printed Crop Top					346	242	286	153	112				1139
ANVI Be Yourself Casual Animal Print Women Purple Top ()					165	222	152	147	77	188			951
Batwing Solid Crop Shirt					364	174	107	135	17	54			851
WOMENS SOLID BLUE ASYMMETRICAL GATHRED SKIRT					140	211	157	94	85	143			830
Dark Blue Raffle Floral Dress					72	159	324	98	75	32			760
Aqua Blue Chiffon Flared Dress					87	112	251	92	100	46			688
Orange Chiffon Flared Dress					96	70	293	37	122	45			663
Tie & Die Shirt Dress					128	153	150	17	86	98			632
Black Chiffon Flared Dress					105	154	142	30	123	57			611
Multi Print Shrit					130	99	41	157	24	144			595
Solid Navy Blue Dress with front Slit					81	117	125	57	88	58			526
Long Maxi Dress					109	129	50	103	24	99			514
Black Buta Top					70	97	133	82	56	45			483
Maroon Velvet Bodycon Slit Dress					107	114	135	44	81				481
Maroon front slit dress					64	111	90	76	50	75			466

Figure 12 : Inventory based on sizes

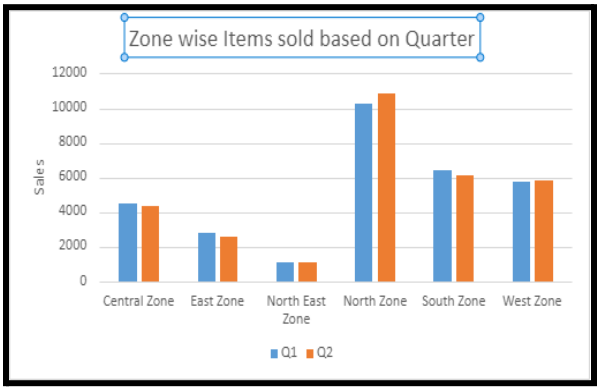
- By doing so, the company will be able to meet the customer demand and also reduce the inventory shelf life leading to lower fabric aging which can also indicate to them that they can introduce a new product into the market. But, a caution to be noted here, if a new product is

released into the market, they should target the high selling zones with minimum inventory and accordingly increase the levels based on the response.

- Inventory can also be controlled by removing old products and replacing them with new ones. But this has to be done with high performing products and only when a new fashion comes into the market.
- This could be done by comparing the sales of the product in the current and previous year and see if there is a declining trend month wise and if that is noted, then that shows a sign that the existing inventory might increase and it is time to bring a change into the market.
- The product life issue could be reduced by identifying the increasing trend quarter wise, by targeting or predicting a native festival that may be coming up which would see an increase in buying power from customers as well and hence trying to sell products during that period.

6 RESULTS AND FINDINGS

- Customer segmentation was the first issue that is being looked at and hence a comparative bar chart is shown below(Figure 13) based on the data table presented above and, to a great extent it can be inferred that North, South and West zones have highest sales in terms of quantity and hence customers should be mainly targeted in these regions but not at the cost of compromising sales in the other zones as well.



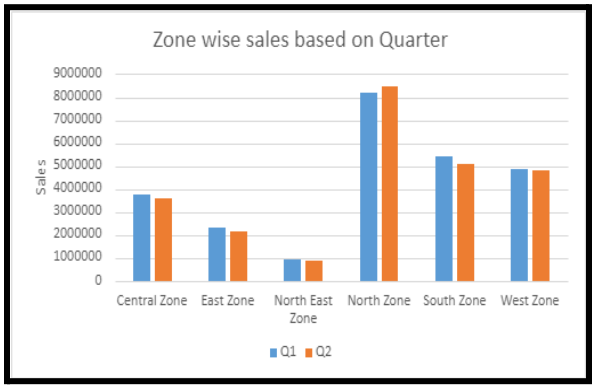


Figure 13 : Zone wise sale(qty) based on quarters *Figure 14 : Zone wise sale(Sell Price) based on quarters*

- This finding could be a doubtful one because some may question that sales(Selling Price) may also be a factor to be taken into account here as quantity sold may not mean that there is an equal amount of revenue earned but figure 14 does strongly support this point.
- Moving towards profitability, as mentioned above that platform wise splitting of sales could be done in order to have the correct marketing strategy and hence improve customer buying and therefore increase sales.
- Below, figure 15, is a plot between an e-commerce platform vs Sales between Jan-June 2023 and here it could be inferred that Myntra is the best performing e-commerce platform and the most consistent as well.
- While all the e-commerce platforms have shown spikes in certain months while during the other months, the sales have come down but, **Nykaa** and **Myntra** have shown a pretty

consistent trend even though their products have been equally distributed in ratio across the platforms.

- Bringing in external factors like festivals also contributes to the increasing demand and sales. As a sharp rise is there in the month of March across most of the platforms, it can be related to the festivals like Holi and many others which leads to huge SALE online which further might have led to this sharp increase in sales for the month of March.
- By this it can be inferred that the company can keep a good amount of inventory by forecasting the festival seasons in India and hence safely assume a good amount of profit in these months.

Figure 15 : Sales across various platforms

- Figure 16 also bolsters the point of more profits and products bought during Q1 when compared to Q2 mainly because of the large number of festivals during those months. 67% of the products were bought during Q1 and 33% during Q2.

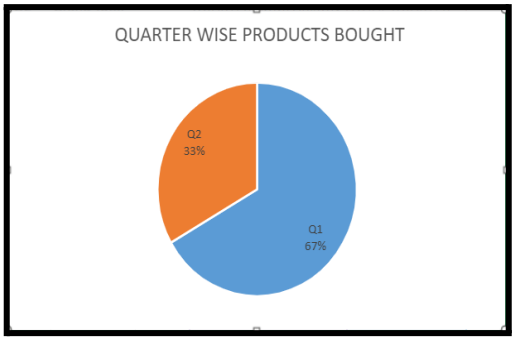


Figure 16 : Quarter wise Products bought

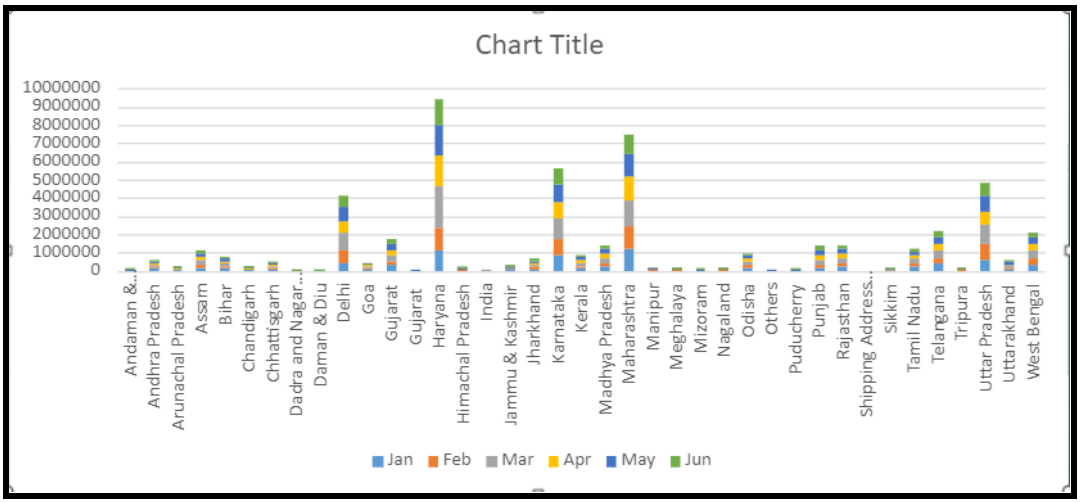


Figure 17 : Monthly sales across different states

- Another parameter that I have looked at is the monthly sales across states(Figure 17) and few of them have outshined the others which can further be used to target customers and which brings the company to keep good inventory levels at the zones that were classified above.
- Diving further into the optimum inventory levels, I have looked at how some of the SKU's(products) are performing in each quarter but I feel that if a better solution is required, a detailed comparison of the performance with the previous year ones needs to be done and then suggest or comment on the optimum inventory level.