E-COMMERCE TRENDS AND PREDICTIONS USING BIG DATA TOOLS

Sreeja Aluguvelly Computer Science University of Missouri, Kansas City Kansas City Missouri USA sa8hm@umkc.edu Roopesh Kesava Reddy Computer Science University of Missouri, Kansas City Kansas City Missouri USA rk4gb@umkc.edu Sneha Muppiri Computer Science University of Missouri, Kansas City Kansas City Missouri USA sm743@umkc.edu Jaswanth Reddy
Narala
Computer Science
University of Missouri
Kansas City
Kansas City Missouri
USA
jnmhc@umkc.edu

1. ABSTRACT

In today's dynamic e-commerce landscape, big data analytics is emerging as a powerful tool for businesses to gain a competitive edge. E-commerce is a strong catalyst for economic development. The rapid growth in usage of Internet and Webbased applications is decreasing operational costs of large enterprises, extending trading opportunities and lowering the financial barriers for active e-commerce participation. Many companies are restructuring their business strategies to attain maximum value in terms of profits as well as customer satisfaction. The objective of this paper is to evaluate the trend and analysis of Flipkart sales. We have used Hadoop for data storage, spark for data analysis and Matplotlib and Seaborn for data visualization.

2. INTRODUCTION

With the increase of e-commerce platforms, vast amounts of data are being generated which can be used to provide key insights for user satisfaction. This paper explores several big data tools and concepts to extract the trends and predictions. The dataset used in this paper is taken from Flipkart, one of the leading e-commerce platforms in India. By using big data tools, patterns in customer purchase behavior, product choices and preferences are analyzed to identify major trends which include seasonal buying patterns and customer sentiment patterns.

This project's primary goal is to exploit data from e-commerce sites, including purchase history, to aid businesses with logistics, profitability optimization, and brand value enhancement through advertising. To enhance business products or services, companies employ data extracted from online product purchase history. These records are monitored to provide insights into customer behavior and market trends.

3. DATA AND METHODOLOGY

We collected our dataset of 4.5GB of data from Kaggle. https://www.kaggle.com/datasets/iyumrahul/flipkartsalesdataset/?select=products.csv. We have two datasets in this project. Sales dataset and Products dataset.

Sales dataset contains the records of all the sales in the month of April in India. It consists of more than 40 million records, and it has 13 columns which include.

- Date
- Order_id,
- Product_id,
- City name,
- Unit selling
- Customer id

and all the required sales columns.

Figure 1 contains the details for the Sales dataset. Figure 2 contains the number of records in the sales dataset.

	Α	1	3	C	D	E	F	G	Н	1	J	K	l	M	N	0
1		Uni	naml	Jnnam date	_	city_name	order_id	cart_id	dim_custom	procured	unit_sellin to	tal_discou	product_id	total_weighte	d_landing	price
2		0	0	0	4/1/2022	Mumbai	112246974	173273802	17995199		1 234	0	344107	202.51303		
3		1	1	1	4/1/2022	Bengaluru	112246976	173273597	18259433		1 64	0	389676	48.714375		
4		2	2	2	4/1/2022	Bengaluru	112247019	173123717	5402601	- 1	1 1031	0	39411	975.996		
5		3	3	3	4/1/2022	HR-NCR	112247045	172547459	15649744		1 57	0	369742	25		
5		4	4	4	4/1/2022	Mumbai	112247123	173081820	10127605		2 30	0	12872	57.980004		
7		5	5	5	4/1/2022	HR-NCR	112247149	173274225	479129		1 65	0	99407	61.5		
3		6	6	6	4/1/2022	Bengaluru	112247170	173273989	8658748		1 115	0	439935	103.05988		
9		7	7	7	4/1/2022	Mumbai	112247182	173273662	16541671		1 20	0	424394	15.399999		
0		8	8	8	4/1/2022	Delhi	112247233	173269263	21466		1 24	0	470636	18		
1		9	9	9	4/1/2022	Delhi	112247239	173274446	13522664	1	1 36	0	311	32.38837		
2	1	.0	10	10	4/1/2022	Delhi	112247317	173274278	13497358		1 20	0	339152	12		
3	- 1	1	11	11	4/1/2022	HR-NCR	112247321	173273059	4480518		1 50	0	18546	46.477013		
4	- 1	2	12	12	4/1/2022	Delhi	112247317	173274278	13497358		1 23	0	4927	14		
5	1	.3	13	13	4/1/2022	Bengaluru	112247328	173269880	18049405		2 19	0	37083	36.1		
6	1	4	14	14	4/1/2022	Delhi	112247393	173273514	5933479		1 130	0	366189	113.04337		
7	1	.5	15	15	4/1/2022	Bengaluru	112247425	173273350	17213227		1 13	0	3928	8		
8	1	6	16	16	4/1/2022	HR-NCR	112247431	173274695	1778240		1 35	0	424982	29.799997		

Figure 1: Sales

Number of rows in the sales: 46706387

Figure 2: Records in Sales

The product dataset contains the data of all the products available for sale in Flipkart, it consists of more than 32000 products, and it has 12 columns which include.

- Product_id
- Product name,
- Product_type,
- Brand name,
- Manufacturer_name

and all the required products columns.

Figure 3 contains the details for the Products dataset. Figure 4 contains the number of records in the Products dataset.

	Α	В	C	D	E	F	G	Н
1		product_id	product_name	unit	product_type	brand_name	manufacturer_	IO_categ
2	0	476763	Christmas - Caro	1 unit	Card		НОТ	Specials
3	1	483436	Plum BodyLovin	20 ml	Sample	Plum BodyLovii	Pureplay Skin S	Specials
4	2	476825	Diwali Gift Card	1 unit	Sample		НОТ	Specials
5	3	483438	Plum BodyLovin	20 ml	Sample	Plum BodyLovii	Pureplay Skin S	Specials
5	4	480473	Flipkart Valentin	1 unit	Card	Flipkart	Dummy Manu	Specials
7	5	483694	Dabur Vita Choo	75 g	Sample	Dabur	Dabur India Pv	Specials
3	6	486016	Plum Green Tea	15 ml	Sample	Plum	Pureplay Skin S	Specials
9	7	486017	Plum Green Tea	15 ml	Sample	Plum	Pureplay Skin S	Specials
0	8	486124	Kari kari Salt & P	22 g	Sample	Kari kari	LT Foods	Specials
1	9	486125	Maggi Liquid Co	180 ml	Sample	Maggi	Nestle India	Specials
2	10	486376	Orion Mango Ch	28 g	Sample	Orion	ORION FOOD \	Specials
3	11	486377	Tata Tea Gold Ca	25 g	Sample	Tata Tea	Tata consumer	Specials
4	12	486559	Conscious Food	50 g	Sample	Conscious Food	Conscious Foo	Specials
5	13	487659	The Laughing Co	15 g	Sample	The Laughing C	Bel Vietnam Co	Specials
6	14	487667	Sunfeast Dark Fa	20 g	Sample	Sunfeast Dark I	ITC Limited	Specials
7	15	487668	Sunfeast All Rou	28.2 g	Sample	Sunfeast	ITC Limited	Specials
8	16	488061	Sleepy Owl Cold	5 x 50 g	Sample	Sleepy Owl	Sleepy Owl Cot	Specials

Figure 3: Products

Number of rows in the products: 32226

Figure 4: Records in Products

4. DATA PRE-PROCESSING

In the dataset we identified there are some columns that are not required for our analysis, we also found that some rows had no meaning at all, and some other rows contained null values. Consequently, we cleaned and pre-processed the data such as removing the unwanted columns and a statistical metric such as the variable's mean has been used in place of missing values.

Figure 4 contains the preprocessed Sales dataset. Figure 5 contains the number of records in the Preprocessed sales dataset.

						+		+	+
202.5130	344107	0.0	234.0	1	17995199	173273802	mbai 112246974	022-04-01	0
48.71437	389676	0.0	64.0	1	18259433	173273597	luru 112246976	022-04-01	1
975.99	39411	0.0	1031.0	1	5402601	173123717	luru 112247019	022-04-01	2
25.	369742	0.0	57.0	1	15649744	172547459	-NCR 112247845	022-04-01	3
57.98000	12872	0.0	30.0	2	10127605	173081820	mbai 112247123	022-04-01	4
61.	99407	0.0	65.0	1	479129	173274225	-NCR 112247149	022-04-01	5
103.05987	439935	0.0	115.0	1	8658748	173273989	luru 112247170	022-04-01	6
15.39999	424394	0.0	20.0	1	16541671	173273662	mbai 112247182	022-04-01	7
18.	470636	0.0	24.0	1	21466	173269263	elhi 112247233	022-04-01	8
32.3883	311	0.0	36.0	1	13522664	173274446	elhi 112247239	022-04-01	9
12.	339152	0.0	20.0	1	13497358	173274278	elhi 112247317	022-04-01	10
46.47703	18546	0.0	50.0	1	4480518	173273059	-NCR 112247321	022-04-01	1
14.	4927	9.9	23.0	1	13497358	173274278	elhi 112247317	022-04-01	12
36.	37083	0.0	19.0	2	18049405	173269880	luru 112247328	022-04-01	13
113.043	366189	0.0	130.0	1	5933479		elhi 112247393	022-04-01	
8.	3928	0.0	13.0	1	17213227	173273350	luru 112247425	022-04-01	15
29.79999	424982	0.0	35.0	1	1778248	173274695	-NCR 112247431	022-04-01	16
85.4198	432764	0.0	76.0	1	7258345	169126739	elhi 112247443	022-04-01	17
150.022	13	0.0	91.0	2	4246317	172994313	-NCR 112247447	022-04-01	18
10.	3913	0.0	16.0	1	13982586	173273818	-NCR 112247464	022-04-01	19

Figure 5: Preprocessed Sales Dataset

Number of rows in the sales: 46627032

Figure 6: Sales Records after Preprocessing

Figure 5 contains the preprocessed Products dataset. Figure 6 contains the number of records in the Preprocessed Products dataset.

_c0 pro	duct_id	product_name	unit	product_type	brand_name	manufacturer_name	10_category	l1_category	10_category_id 11_	categor
1	483436	Plum BodyLovin' H	20 ml	Sample	Plum BodyLovin'	Pureplay Skin Sci	Specials	Free Store	343	
3	483438	Plum BodyLovin' T	20 ml	Sample	Plum BodyLovin'	Pureplay Skin Sci	Specials	Free Store	343	
4	480473	Flipkart Valentin	1 unit	Card	Flipkart	Dummy Manufacturer	Specials	Bill Breaker	343	
5	483694	Dabur Vita Chocol	75 g	Sample	Dabur	Dabur India Pvt Ltd	Specials	Free Store	343	
6	486016	Plum Green Tea Po	15 ml	Sample	Plum	Pureplay Skin Sci	Specials	Free Store	343	
7	486017	Plum Green Tea Oi	15 ml	Sample	Plum	Pureplay Skin Sci	Specials	Free Store	343	
8	486124	Kari kari Salt &	22 g	Sample	Kari kari	LT Foods	Specials	Freebie	343	
9	486125	Maggi Liquid Coco	180 ml	Sample	Maggi	Nestle India	Specials	Freebie	343	
10	486376	Orion Mango Choco	28 g	Sample	Orion	ORION FOOD VINA C	Specials	Freebie	343	
11	486377	Tata Tea Gold Car	25 g	Sample	Tata Tea	Tata consumer pro	Specials	Freebie	343	
12	486559	Conscious Food Or	50 g	Sample	Conscious Food	Conscious Food	Specials	Freebie	343	
13	487659	The Laughing Cow	15 g	Sample	The Laughing Cow	Bel Vietnam Co.,	Specials	Freebie	343	
14	487667	Sunfeast Dark Fan	20 g	Sample	Sunfeast Dark Fan	ITC Limited	Specials	Freebie	343	
15	487668	Sunfeast All Roun	28.2 g	Sample	Sunfeast	ITC Limited	Specials	Freebie	343	
16	488061	Sleepy Owl Cold B	5 x 50 g	Sample	Sleepy Owl	Sleepy Owl Coffee	Specials	Freebie	343	
17	488062	Sleepy Owl Cold B	5 x 50 g	Sample	Sleepy Owl	Sleepy Owl Coffee	Specials	Freebie	343	
18	488963	Sleepy Owl Cold B	5 x 50 g	Sample	Sleepy Owl	Sleepy Owl Coffee	Specials	Freebie	343	
19	488059	Sleepy Owl Cinnam 1	pack (5 x 50 g)	Sample	Sleepy Owl	Sleepy Owl Coffee	Specials	Freebie	343	
20	488058	Sleepy Owl Carame	5 x 50 g	Sample	Sleepy Owl	Sleepy Owl Coffee	Specials	Freebie	343	
21	488060	Sleepy Owl Cinnam	3 x 50 g	Sample		Sleepy Owl Coffee		Freebie	343	

Figure 7: Preprocessed Products Dataset

Number of rows in the DataFrame: 29030

Figure 8: Products Records after Preprocessing

5. Data Analysis

We started a spark session which connects to HDFS and reads data from the data frame which we loaded in the Flipkart directory in HDFS. Using this data frame, we started performing analysis.

Figure 9 shows the analysis of orders placed in each city. where we plotted the graph for it.

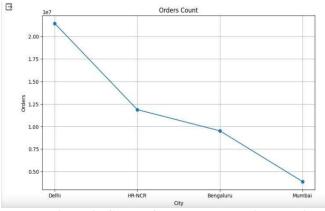


Figure 9: Count of orders recorded in each city.

Figure 10 shows the analysis of product categories where we plotted the graph of the top 10 product categories which recorded more orders.

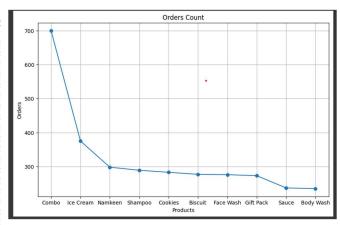


Figure 10: Top 10 Product Categories which recorded more orders.

These are some of our preliminary analyses. With these preliminary analyses we started analyzing sales based on multi categories like date-city, city-month, and many more.

In our data analysis, at first, we started with the total sales happening in all the four cities namely Mumbai, Delhi, Bengaluru, and HR-NCR available in the data set for April, May, June and July months.

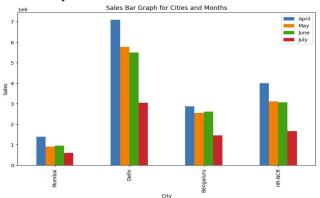


Figure 11: Sales recorded in each City in each Month.

From the sales bar graph for cities and months, we can obtain a few insights.

- Delhi has recorded a higher number of sales than the other three cities in all the four months.
- Sales are higher in April than in the other three months in all the cities.

This might be because of the start of summer holiday season in India. This observation helps the e-commerce companies to plan their sales well ahead and come up with attractive marketing strategies to attract more people and they can even stock up more things during this time so that they can maintain customer satisfaction well.

While coming to the analysis of average daily sales in all categories in each city.

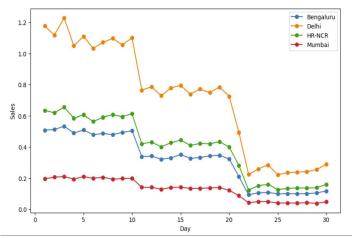


Figure 12: Daily sales in millions in each City.

Here we can observe an interesting pattern that maximum sales in all the cities are happening only in the first 10 days of the month and after that we can observe a steep decrease in the daily sales and towards the end of the month, there were little ups and downs in the pattern. This shows an interesting observation that people are more likely to purchase an item at the beginning of the month rather than during the middle of the month. This analysis will give an idea of the people's purchase patterns and they can increase the procured items at a particular city warehouse so that they can deliver products to the customers in a faster and secure manner.

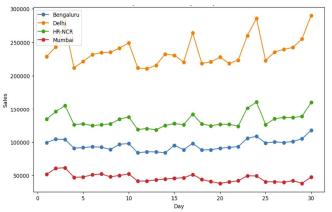


Figure 13: Daily sales in April Month in each City.

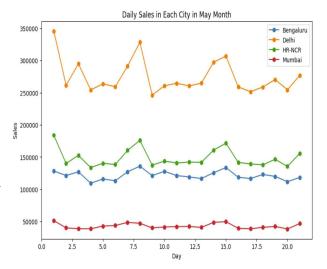


Figure 14: Daily sales in May month in each City.

When we break down the trends of the April and May month, In April month daily sales we can see contradicted patterns than what we have discussed above. The daily sales in all the cities have begun to rise more during mid-April than at the start of the month. From our observation, we found that it might be because of the attractive discounts e-commerce platforms have offered to make use of the holiday season. Coming to the May month, our starting trend continues to follow. It is a challenging task to describe people's purchase patterns without giving any particular details about discounted sales and the special holiday seasons.

We have obtained a list of several products ranging from grocery to cleaning essentials which were sold more in all the cities. From that list we have selected to see the trends of Onion daily sales in April month as this is the most used cooking product in India.

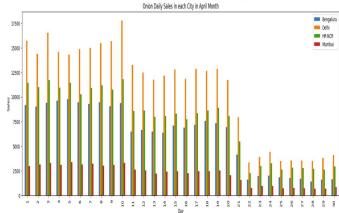


Figure 15: Onion April Daily sales in each City.

In Onion daily sales, Delhi recorded the most and Mumbai the least. Also, the sales were higher in the first 20 days of the month and later there was a gradual decrease in the sales towards the month-end. This analysis helps the business vendors to procure more stock during the initial days of the month so that the fresh products will attract customers more and which in turn maximize the profits.

The top 10 brands sold across all the cities are GHH, GMC, Aplus, GHD, Amul, Town Grocer, Haldiram's, Apple, SaveMore, Grocery.

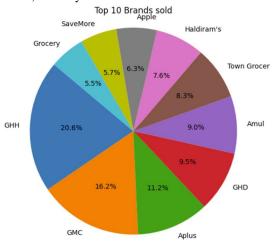


Figure 16: Top 10 Brands Sold.

Most of the brands from the list are related to household items which are used in everyday cooking. This pie-chart also tells how people are showing interest in purchasing products for their daily use from e-commerce platforms.

Post COVID-19 season, we identified that people are keener in maintaining personal hygiene and household hygiene. So we tried to analyze how the sales of cleaning essentials products are happening.

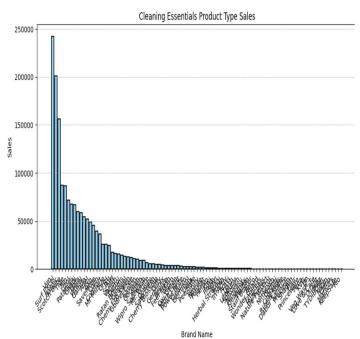


Figure 17: Cleaning Essentials Brands sales.

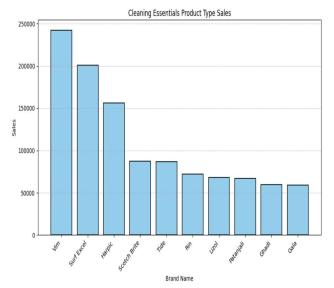


Figure 18: Top 10 Cleaning Essentials Brand sales.

We obtained a clumsy graph as there are numerous brands, so we cut down our analysis into top 10 cleaning essential product brands.

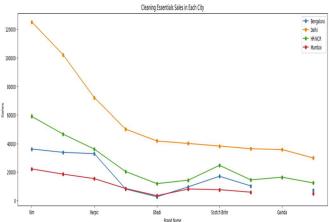


Figure 19: Cleaning Essentials Brand sales in each City.

Even in the sales of cleaning essentials, Delhi has recorded more and Mumbai the least.

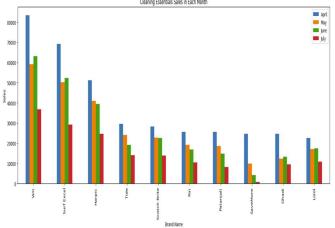


Figure 20: Cleaning Essentials Brand sales in each Month.

Sales in April are more when compared to the other months. This also might have an influence on the holiday which we have discussed earlier.

6. Server Client Architecture

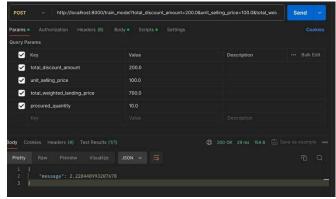


Figure 21: Client (Postman)

	Wait	ting for	r applio	cation	startu	D.					
						μ.					
	App.	TCat10	n startı	пр сошр	Tere.						
FO:	llvir	orn rui	nnina or	httn:	//0.0.	A.A:RARA (Press CTRL+C t	n muit)			
	V121	.0111 101	inizing of		,,0.0.	0.010000 (TOUS STREET	o quit,			
of rows	in the sale	51 46706357									
10 22:30:	29 KARN CST	HeaderChark	er: CSV head	er doss net	conform t	to the schene.					id, total_weighted_landing_p
n: _co, U	nnamed: 0.1	, Unnamed I	d, date_, ci	ty_rate, er	der_is, co	rt_id, dia_custos	rer_key, protutes_quant	ity, unit_telli	g_price, total_discan	nt_amount, prod.	et_id, total_weighted_landis
	ut found: ///Users/on	cha/Download	ds/archiveNZ	91201/Sales							
	.1 Unnancd:										_weighted_landing_price
	*1	915922-94-1	91 Martina i	11122469741	179279882	12995199	11	234.81	9.4	9441971	202.51363
		1 2022-04-4	eljeccoluru	[122246976]	178273097	18209433			0.4	389676	
		2 2822-86-4	01 Bergeluru 01 HR-NCR	112247619	2231237257 2230777550	5462681 15649744		1831.8	9.8	39411 369742	975.996
		5 2022-86-	Ri Parkei	113247123	1730161828	18127685		38.8	9.8	12872	57,998861
		512022-04-4	esi He-NCE	11122471491:	178274225	479129		66.81	9.81	99487	
		6 2022-86-	81 Sengaluru	1112247178	173273989	8658748 18541671		115.8 20.8	9.4	439935 424394	183.859875
	4	8 2022-04-4	er Homour	122247233				24.8	0.4	478686	15.399999
		9 2022-04-4	etj Galbi	112247239	272274446	13522464		26.8	9.8		32.38837
		10 2022-04-1		122247817				20.0	0.0		22.0
	11	11 2022-06-4	et He-ACE	112247321	17327365V	6488118 13497368		10.6 22.0	9.8	18546 4927	46.477913
			01 Bengeluru	112247328	173269838	18049403		19.01	6.4	37883	26.1
	16	14 2022 06-1	241ed 10	111247191	172271614	5923479		110.0	9.4	364589	313.04327
		15 2022-06-4	ei Bengeluru	1122247426	270273050	17219227 1779240		10.0] 35.0]	0.0	0929 424982	20,799997
		17 2922-96-4	24[a0	112247448	140174710	7259346		76.8	1.1	432764	86,42983
	181	18 7022-06-4	eti He-NCE	112247647	172994313	6246217		91.8	0.0		158.82224
	19	19 2022-04-1	ES HR-NOR	1112247464	178273818	19982586		16.8	9.4		59.9
owing to	p 10 rows										
	oftw man	e. erder id	cart id d	in customer	key area	po the schema.	nit selling grice, teta	discount arms	or product id. total	welchted landing	
8 22:38:							unit selling price, to				
: , date											
: , date : _c0, d :d: _c0 b	ut found:										
: , date : _c0, d d: _c0 b e: file:	ut found: ///Users/sm										
: , date : _c0, d d: _c0 b e: file: date_	ut found: ///Users/sn city_name	order_id	cert_id die	_oustener_k	sylprocure		selling_price total_ele	count_amount pp	odust_id total_weighte		
: , date : _c0, d d: _c0 b e: file: date_	ut found: ///Users/sn city_name Mumbail	order_id	cert_id die	_oustemar_k	sylprocure		294.0	0.0]	944197	202.51303	
: , date : _c0, d :d: _c0 b !e: file: date_ :22-84-01	ut found: ///Users/sm city_namo Wumbai Bengaluru	order_id 112245974 1 112245976 1	cert_1d die 73273862 73273597	_oustener_k	sy procure 99 33		234.0	0.01 0.0		202.51303 48.714375	
date_ date_ date_ file:	ut found: ///Users/sm city_name Mumbul Bengaluru Bengaluru	order_id 112245974 1 112245976 1 112247919 1 112247945 1	oert_id die 73275862 73275697 73123717 72547469	_oustemar_k 170951 182504 64020 156407	99 procure 99 33 95 66		234.0 64.0 3631.0 67.0	9.0 9.0 9.0	944197 389676 39413 369742	202.51503 45.714375 975.996 25.0	
date_ date_ date_ 22 84 01 122 84 01 122 84 01	ut found: ///Users/sm Dity_name Pumbul Bengaluru Bengaluru BR-ACR	order_id 112245974 1 112245976 1 112247919 1 112247919 1	oart_id die 73275862 73275597 731237377 73123737 73547469 73853828	_Dustewar_k 179961 182594 64926 156497 182276	89 process 99 33 95 64 96		234.0 64.0 383.0 67.0 38.0	0.01 0.01 0.01 0.01 0.01	344187 389676 39413 369742 32872	202.51303 45.714375 976.996 20.0 57.980084	
date_ date_ date_ file: file: date_ file: file: date_ file: file: date_ file:	ut found: ///Users/sm city_name Fusball Bengaluru Bengaluru Benkell Fusball Fusball Fusball	ordor_id 112240974 1 112240976 1 112240976 1 112247945 1 112247945 1 112247149 1	0ert_1d die 73275802 73275897 73123717 73123717 73181328 73274225	_Dustener_k 179951 182564 54420 154467 181270 4761	ey procure 99 33 96 64 98 29		234.0 64.0 1831.0 67.0 38.0 65.0	0.0] 0.0] 0.0] 0.0] 0.0]	944187 359476 39411 369742 13872 99487	202.51203 48.714376 975.996 25.9 57.98684 41.5	
date co b les files file	ut found: ///Users/on City_name Eury_name Mumbal Bengaluru He-hCS Mumbal He-hCS Mumbal He-hCS Mumbal	ordor_id 112240974 1 112240976 1 112240910 1 112247985 1 112247149 1 112247149 1 112247149 1	cert_id die 73275802 73275597 7312371/ 72547469 73801628 73275909 73275962	_oustener_k 179981 182584 54820 154427 201270 4791 86587 162416	8y precure 99 33 95 64 96 29 64 73		234.0 64.0 3831.0 67.0 38.0 45.0 235.0 28.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	944197 389676 39913 369742 13872 90467 439935 449994	202.51383 48.714375 975.996 25.9 87.98686 61.5 183.85675 15.395995	
11 , date 12 c0 , d 13 c0 , d 14 c0 b 14 c1 file: 12 c4 c1 12 c4 c1 12 c4 c1 12 c5 c5 12 c	ut found: ///Users/en city_name Musbul Bengaluru Bengaluru Musbul HP-hCR Bengaluru Wusbul Delbi	order_id 112249974 1 112249974 1 112249919 1 112247919 1 112247149 1 112247149 1 112247149 1 112247149 1	cert_id die 73275802 73273597 7323717 7323717 73254769 7383529 73274225 73275862 73275862 73275862	_oustemar_k 179951 18254 54826 156407 181276 4791 865476 214	py procure py 33 95 64 99 29 66		234.0 64.0 1031.0 67.0 38.0 65.0 115.0 24.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	344187 359676 39943 369742 249742 23872 90487 43994 43994 479436	202.51303 48.714375 976.996 25.9 87.98084 61.5 183.890875 15.39999 18.8	
date :	ut found: ///Users/en city_name Mumbal Bengaluru BenACR Mumbal Me-ACR Mumbal Bengaluru Mumbal Bengaluru Bengaluru Delhi	order_1d] 112249974[1 112749974[1 112749976]1 112747946[1 112747946]1 112747179[1 112747179]1 112747723[1 112747723]1	cart_1d die 73273882 73273597 73329717 73547459 73873499 73274295 73273999 73274999 73274999 73274999 73274999 73274446	_oustener_k 179981 182594 54826 154497 201270 4791 86587 165416 214 130276	sy procure 991 331 961 441 991 291 481 71 661 944		234.8 64.6 383.8 67.0 38.8 65.8 115.8 24.8 24.8	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	944187 389076 39911 39911 367702 12872 99467 43993 43994 479036 311	202.51303 48.714375 978.998 25.8 97.93084 61.5 183.859875 15.39999 18.8 32.38637	
date :	ut found: ///Users/en city_name Kumbul Bengaluru Bengaluru Wumbul He-NCS Bengaluru Yumbul Delei Delei	order_id 112249974 1 112249974 1 112249919 1 112247919 1 112247149 1 112247149 1 112247149 1 112247149 1	cort_1d die 73273802 73273897 7323397 7313397 73133828 73274295 73273808 73273808 73273808 73274278 73274278	_oustemar_k 179951 18254 54826 156407 181276 4791 865476 214	sy procure 991 33 35 36 44 38 29 48 71 46 58		234.0 64.0 1031.0 67.0 38.0 65.0 115.0 24.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	344187 359676 39943 369742 249742 23872 90487 43994 43994 479436	202.51303 48.714375 976.996 25.9 87.98084 61.5 183.890875 15.39999 18.8	
date da	ut found: ///Users/on Gity_mare Murbul Bengeluru Bengeluru Murbul Murbul Murbul Beleli Deleli Deleli Beleli Deleli Beleli Deleli Beleli Deleli	order_id 112249974 1 112749974 1 112749976 1 112747985 1 112747785 1 112747783 1 112747783 1 112747783 1 112747783 1 112747783 1 112747787 1 112747787 1	cart_id die 73273862; 73273867; 73123717; 73123717; 73123717; 7327369; 7327369; 7327369; 7327428; 7327428; 7327428; 7327428; 7327428; 7327428;	_customar_k 179941 187544 54427 184276 4791 185877 165416 214 1312276 131473 44885 151473	sy procure 991 331 901 641 291 661 561 561 561 561 561		234.0 64.0 1001.0 67.0 38.0 65.0 115.0 28.0 24.0 30.0 28.0 28.0 28.0 28.0 28.0 28.0 28.0 2	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	344187 339476 39411 369742 11872 99487 43998 439994 439994 479036 311 339152 15946 4927	202.51393 45.714375 975.993 25.8) 57.906064 64.5 181.850975 18.9 92.38637 12.8) 46.477813 14.8)	
date co. 6	ut found: ///Users/on ///Users/on ///Users/on ///Users/on ///Users/on ////Users/on ///////////////////////////////////	order_id 112349974 1 112349974 1 112349974 1 1123499991 1123499991 11234971991 11234971991 1123497391 1123497391 1123497391 1123497391 1123497391	oart_id die 79273862; 79273897 79129797 79129797 79129797 7927469 79274625 79273662 79273662 79273662 79273662 7927367 7927367 7927367 7927367 7927367 7927367	_oustoner_k 179961 183064 64829 154467 184276 4791 86587 16916 214 135226 134973 44880 134973 154973	py procure py procure py		234.0 64.0 551.0 57.0 38.0 65.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 2	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	944197 399413 99413 99413 90740 12872 90497 439075 439075 439074 478436 311 339152 13536 4927 27893	202, 81303 48,714375 978, 998 57, 918004 64, 5 183, 959975 18, 932, 38337 18, 6 32, 38337 14, 6 45, 477813 14, 6 36, 1	
date date date date date date date date	ut feund: ///Users/on ///Users/on (city_mame) Mumbul Bengaluru Bengaluru Mumbul He-NCB Bengaluru Bubbi Delhi Bengaluru	order_id 112249974 1 112249974 1 112269191 112267869 1 1122671869 1 112267189 1 112267189 1 112267289 1 112267289 1 112267289 1 112267319 1 112267319 1 112267319 1	cart_id die 79279862; 73279807; 73129712; 73129712; 73129712; 73274295; 73274295; 73274293; 73274293; 73274293; 73274273; 73274273; 73274273; 73274273; 73274273; 73274273; 73274273; 73274273; 73274273; 73274273; 73274273;	_customar_k 179961 182084 84926 154477 182276 4791 169416 214 13:220 13:4773 44880 15:4973 189494 59334	sy procure 331 80 64 81 29 66 55 56 331 58 58 331 58 58 331 58		234.0 64.0 1001.0 57.0 38.0 65.0 115.0 24.0 34.0 24.0 35.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 2	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	944197 395075 97413 97413 367702 50677 405905 405905 405905 405905 313 397602 389602 389602 389602 389602 389602 389602 389602 389602 389602 389602 389602 389602 389602	202.51393 48.714376 978.998 20.8) 97.998894 64.5 123.89999 18.8 32.38937 12.8 46.477813 14.8 36.1 13.84337	
a: c0, d ad: c2 h ad: c2 h ad: c3 h ad: c4 h ad: c5 h ad:	ut found: ///Users/sn city_name; city_name; city_name; lengaluru;	order_id 112349974 1 112349974 1 112349974 1 1123499991 1123499991 11234971991 11234971991 1123497391 1123497391 1123497391 1123497391 1123497391	cort_10 die 732738621 73273597 73123121 / 73123121 / 73274295 73274295 73273662 73273662 73274406 73274369 73274369 73274369 73274369 73274369 73274369 73274369 73274369 73274369	_oustoner_k 179961 183064 64829 154467 184276 4791 86587 16916 214 135226 134973 44880 134973 154973	sy procure		234.0 64.0 551.0 57.0 38.0 65.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 2	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	944197 399413 99413 99413 90740 12872 90497 439075 439075 439074 478436 311 339152 13536 4927 27893	202, 81303 48,714375 978, 998 57, 918004 64, 5 183, 959975 18, 932, 38337 18, 6 32, 38337 14, 6 45, 477813 14, 6 36, 1	

Figure 22: Server (Local Host)

7. CONCLUSION

The report represents a comprehensive approach of E-commerce trends data analysis and trends using big data tools, The proposed methodology was applied to the transaction data acquired from Kaggle, which included approximately 467,06,387 records of users and the items they purchased, viewed with their respective location.

Initially we performed data cleaning and then pre-processing it using spark. We have come up with some outcomes like the total sales occurred in all the cities with respective consecutive months. We also come across the analysis of average daily sales of all kinds of products in each city, while we break down the trends of sales of products during each month and get to know that maximum sales are going to happen in the first ten days and the top 10 types of products which are going to sell more. Effective dynamic charts and graphs were made to show the

outcomes, which helped other users understand them clearly about all the results and findings.

Based on our research, the recommended method provides a solid and scalable means for businesses to enhance their business and optimize their operations by gaining insights from e-commerce data. This can lead to better business outcomes. Businesses can use this approach to analyze e-commerce data and gain valuable insights into consumer behavior, market dynamics, and trends by harnessing the power of big data. By personalizing the customer experience, companies can increase customer satisfaction and loyalty with dynamic pricing strategies, customized recommendations, and focused advertising. When companies have a better grasp of their customers and the market than their rivals, they may develop more effective marketing campaigns and product offers.