

# Sales Analysis of apparel brand

## Table of contents:

1. [Context](#)
2. [Dataset](#)
3. [Features of dataset](#)
4. [Assumptions](#)
5. [Basic Data cleaning](#)
6. [Create additional fields](#)
7. [Break down of Analysis](#)
8. [Data Processing](#)
9. [Recommendations](#)
10. [Tableau Data Visualization](#) - Links

## Context:

Leading AI and analytics company helping some of the world's leading brands solve their complex data challenges. Their end-to-end platform enables companies to leverage the power of data to craft strategies, create engaging customer experiences, and drive measurable business impact. They were voted one of the best CRM solutions providers in India by the CICC Review magazine.

## Dataset: [Link](#)

This dataset has information on more than 500k transactions from 2020 to 2022 at brand A across different stores in UAE and Qatar.

## Features of dataset:

Features	Description
customer_id	unique identifier for a customer for a given transaction.

current Tier	is a way of categorizing customers based on purchase history
customer_nationality	nationality of the customer
date_transacted	date of transaction
storeId	store from where the transaction has happened
store_city	city of where the store is located
transactionId	Way of identifying a particular transaction.
ItemId	Id of item
Brand	A (name of the brand)
ItemName	name of the item.
Product_category	category under which the product falls.
High_level_product_category	broader categories of product_category.
quantity	total quantity of the item the customer purchased.

total_spend	total amount the customer spent on that item within that transaction.
signed_up_loyalty_program_date	date that the customer signed up for a loyalty program.
signed_up_app_date	date that the customer signed up on the app.

### Assumptions:

1. There are negative values in the total spend column, which we assume represents a refund given by the brand.
2. There are negative values in the quantity column, which we assume that the customer returned the item.
3. The Total spend column contains values of zero, which we assume represent a free item(freebie) given to the customer by the brand.

### Basic Data cleaning:

1. Add a data [source filter](#) to remove null transaction ids.
2. [Change the datatype](#) of Itemid field from number to string.
3. Assign geographic role country/region to customer nationality field.
4. Convert Item id to dimension.

### Create additional fields that will make analyzing the data easier in Tableau

1. Create a field called **Registered or not** to check whether a customer is registered or not.
  - o A customer is considered to be registered only when the customer id field is not null and signed up loyalty program date field is also not null
  - o Create a calculated field->name it **Registered or not**
  - o Use formula

```
IF (NOT ISNULL([Customer Id])) AND (NOT ISNULL([Signed Up Loyalty  
Program Date]))  
THEN  
"Registered"  
ELSE  
"Not Registered"  
END
```

- We are checking signed up loyalty program date fields also because in the dataset we have some customers where customer id is present but we do not have their signed up date.

2. Create a field called ***Refund transaction or not*** to determine whether or not the transaction was a refund

- Based on the assumption that if the total spend field is less than 0 then it was a refund given to the customer and if the total spend was greater than 0 it was a normal sales transaction and if the total spend value was 0 and quantity was greater than 0 then it was a freebie given to the customer.
- Create a calculated field and name it **Refund transaction or not**
- Use formula-

```
IF [Total Spend]<0  
THEN  
"Refund transaction"  
ELSEIF [Total Spend]=0  
THEN  
"Freebies"  
ELSE  
"Sales"  
END
```

3. Create a field called ***Item refunded or not*** to determine whether or not the item was refunded.

- We are checking whether the quantity field was less than 0 or not, if the quantity is less than 0 then refund else not a refund
- Create a calculated field and name it **Item refunded or not**

- Use formula

IIF([Quantity]<0,"Item refunded","Item not refunded")

**Note:** You would have the following Distinct values after creating the above three calculated fields or when you drag any of the fields either on rows or column shelf you would see following Distinct values:

- Registered or not
    - Registered
    - Not Registered
  - Refund transaction or not
    - Refund
    - Freebies
    - Sale
  - Item refunded or not
    - Item refund
    - Item not refund
4. Similarly created additional fields like First transaction date, Last transaction date, Item Refunded or not, Pre\_Loyalty\_Transactions, Retention, Retention rate, Post\_Loyalty\_Transactions, Refund transaction or not, Registered or not.
  5. For the **current tier** field edit the **Null** entry with an alias **No current tier**.
  6. For the **customer nationality** field edit the **Null** entry with an alias **No nationality**.

## Break down of Analysis

1. Customer Behavior and Demographics
  - A. Customer Segmentation
  - B. Loyalty Program Analysis
  - C. Customer Lifetime Value (CLV)
2. Sales and Revenue Analysis
  - A. Revenue Breakdown
  - B. Product Performance
  - C. Sales Trends and Patterns
3. Operational Insights
  - A. Store Performance
4. Comparative Analysis

- A. Cross-Store Comparisons
- B. Product comparisons

## Data Processing:

### Customer Behavior and Demographics

#### A. Customer Segmentation:

How do purchase behaviors differ among customers from different nationalities?

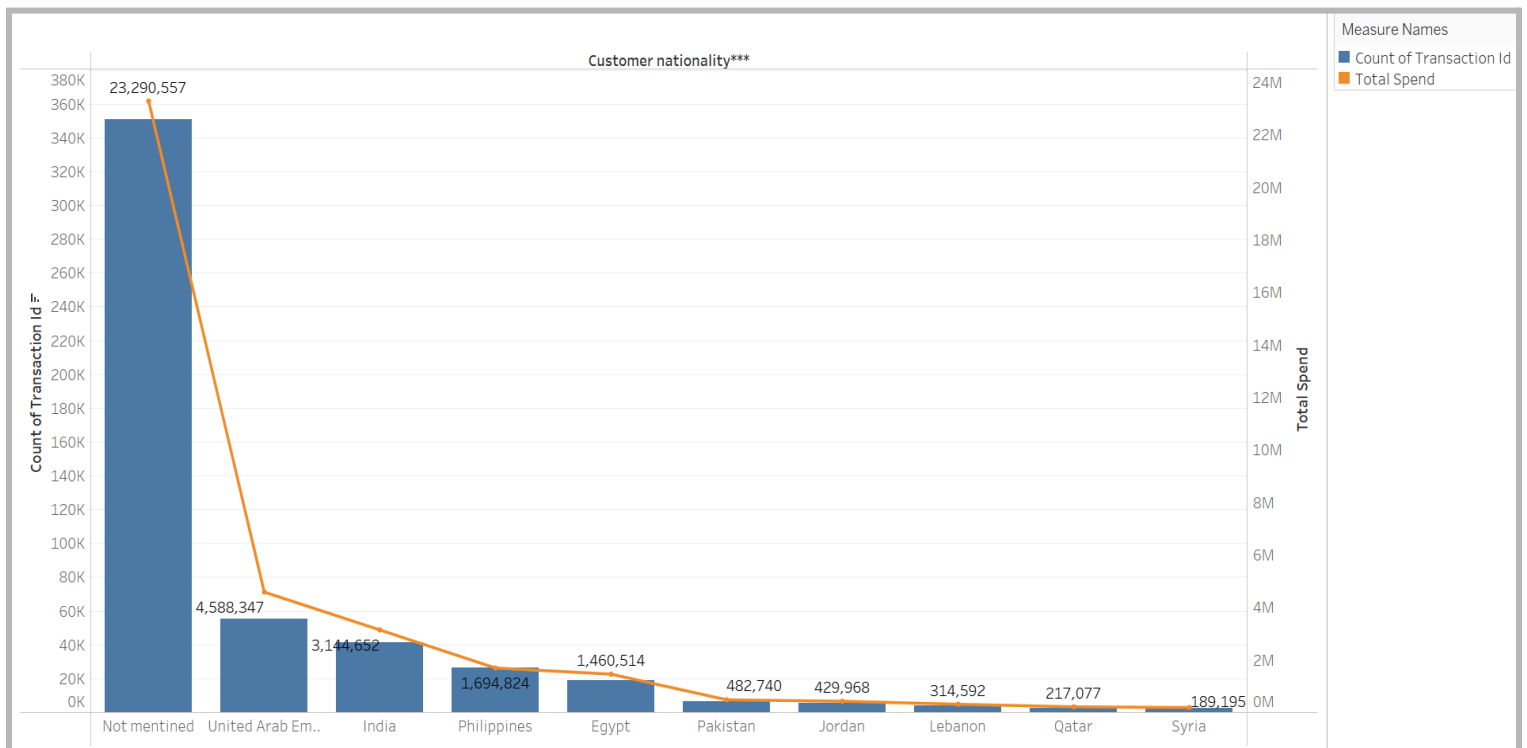


Fig shows Count of transaction & total spend of different customer nationality

**Insight:** Top 5 countries customers are from are UAE, India, Philippines, Egypt, Pakistan. The sales volume & total spend also follow similar order. UAE customers made 55k transactions, 41k by Indians & 26k by Filipinos. Whereas, 46 lakh \$ by UAE, 31 lakh \$ by Indians & 17 lakh \$ by Filipinos.

What are the spending patterns of customers in different customer tiers?

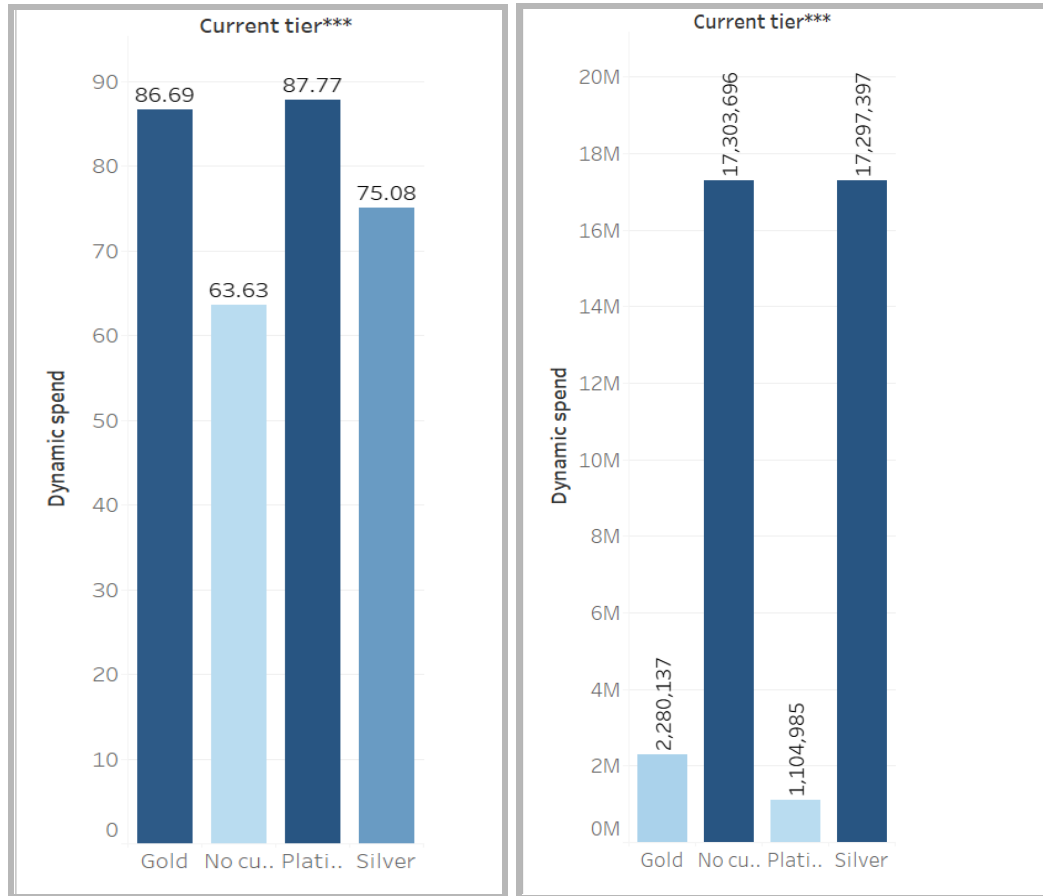


Fig shows spending patterns of customers of different segments in terms of average spend & total spend.

#### Current tier.. ₹

Silver	62,386
Gold	2,073
Platinum	488
No current tier	42

In Fig color represents numbers of customers.

**Insight:** There are 3 customer segments Silver, Gold & Platinum.

Silver segment i.e, 62,386 people have average spend of 75 & 1.72cr total sales.

Gold with 2073 people having average 86 & total of 22 lakh.

Platinum with minimum of 488 customers had highest avg of 87.77 & 11 lakh sales.

And platinum has the highest sale per customer of 2254\$.

How will the average order value change over time wrt to store city ?

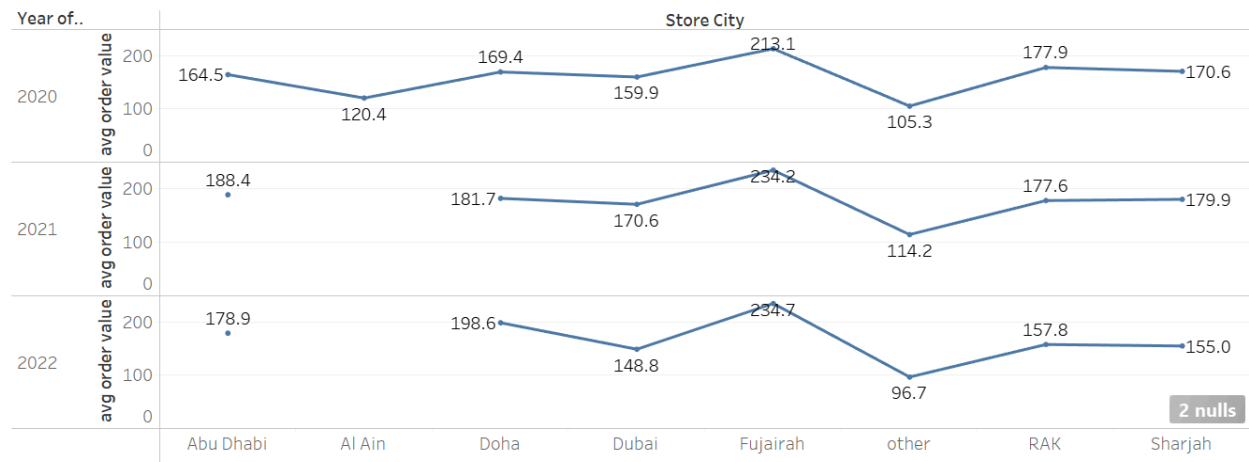


Fig tells change of AOV across store city in different timeline.

**Insight:** The AOV is highest in Fujairah city all 3 years with 234\$. Al Ain has least Aov of 120\$. Rest all cities with Aov ranges between 160-188\$.

How does the transaction frequency vary by customer segment?

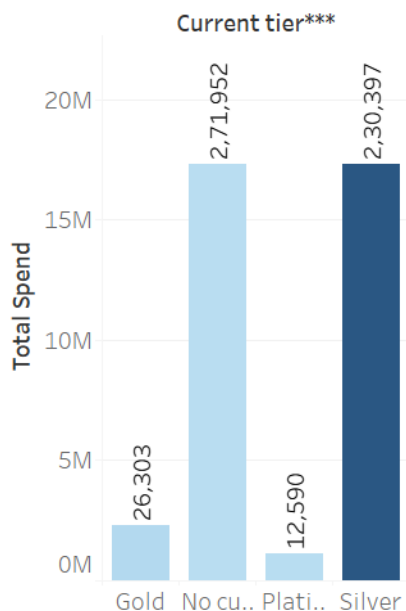


Fig tells us transaction frequency among different customer segments. Color indicates the number of customers.



**Insight:** As the Silver with the highest customers has the highest sales volume of 2.30lakh transactions. Gold with 2k customers made 26k transactions. 500 platinum customers made 12500 transactions. Tells Number of customers is directly proportional to Transaction frequency.

1. Loyalty Program Analysis:

What is the impact of signing up for the loyalty program on total spend and transaction frequency?

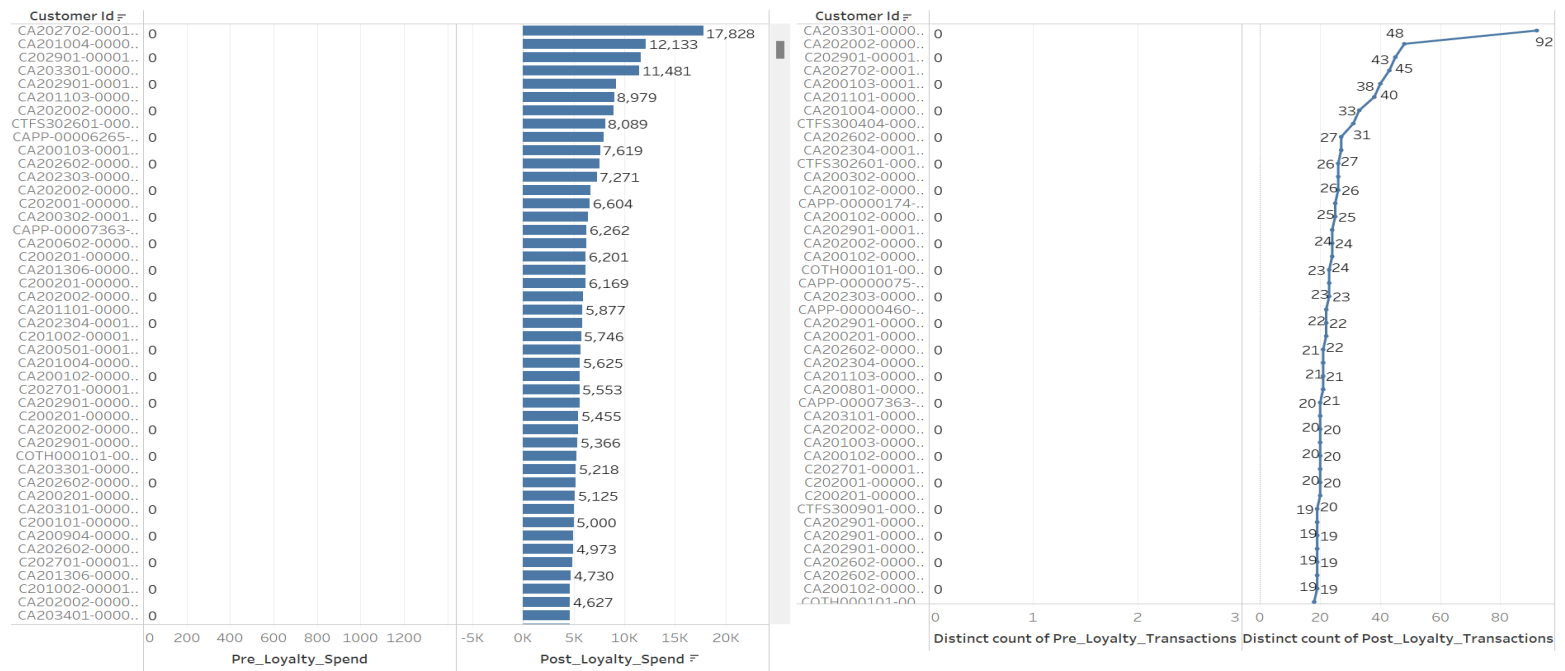


Fig shows pattern after loyalty program in total spend & transaction frequency

**Insight:** 99.5% of the pattern tells that customer made the first purchase after loyalty program only.

How many customers signed up for the loyalty program over time?

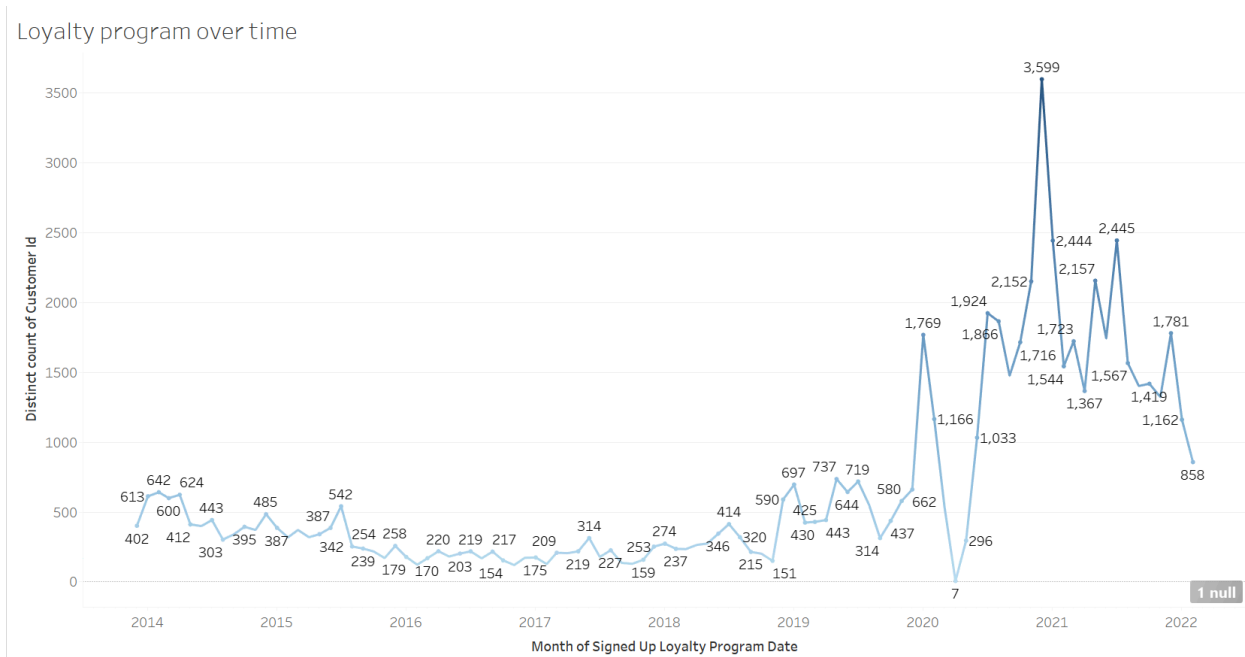


Fig represents Trend of loyalty program sign ups wrt No. of customer

**Insight:** From 2014 & 2015 had avg loyalty signup of 400-500/month. From 2016 till 2019 the avg fell to 200 signups/month. Then 2019 avg increased to 500-700 signup/month. In April 2020 had lowest signup of 7 & in December 2020 had 3600 signup/month. After 2020 avg of above 1500 was maintained.

## 2. Customer Lifetime Value (CLV):

What is the average CLV for customers in different tiers?

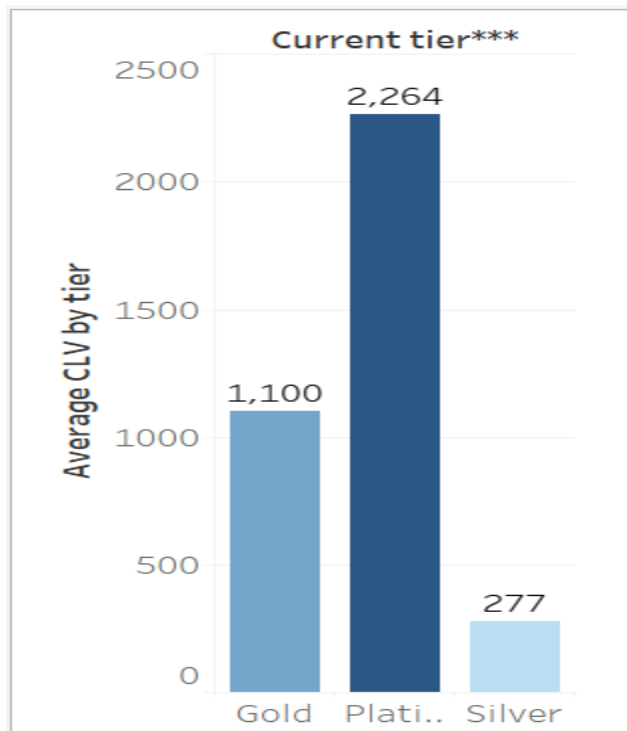


Fig tells Avg CLV by tier wrt different customer segment

**Insight:** Average CLV of Platinum customers is highest of 2200\$ and silver has least avg of 277\$

How does the CLV vary across different customer nationalities?

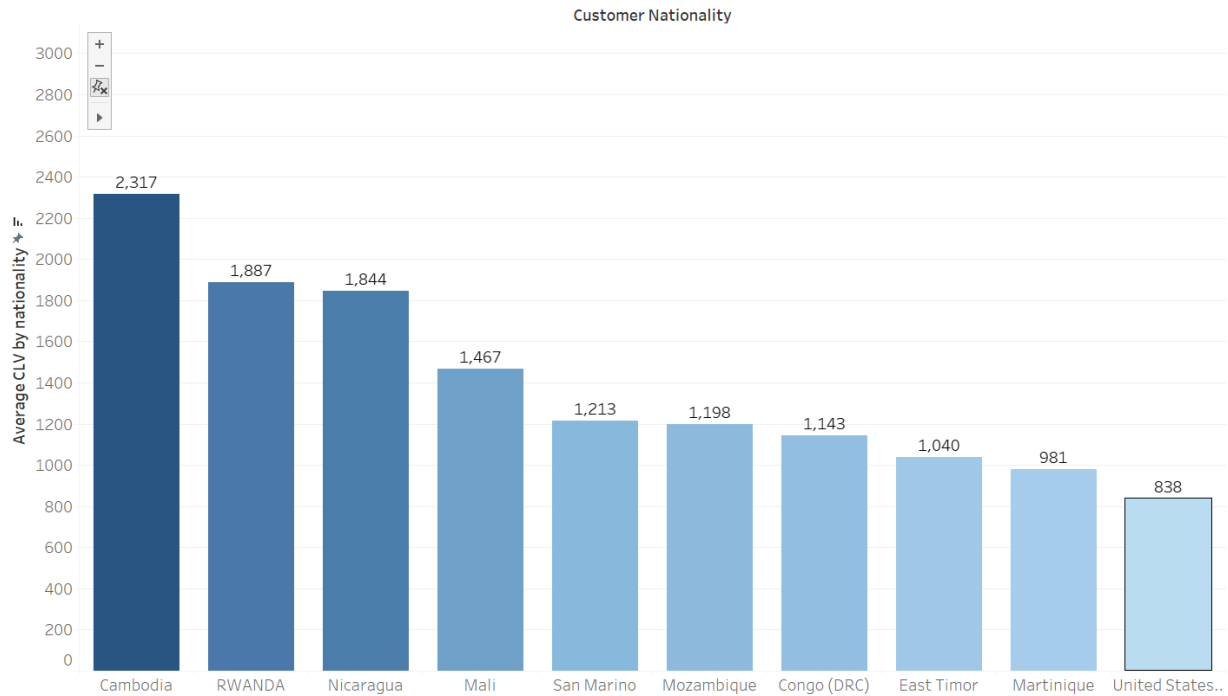


Fig represents Avg CLV by nationality wrt to cust nationality

**Insight:** CLV is highest in Cambodia with 2317\$. Rwanda & Nicaragua with 1800\$ rest customers from countries Mali, San Marino, Mozambique, Congo & East timor have purchased above 1000\$.

## Sales and Revenue Analysis

### 1. Revenue Breakdown:

What is the total revenue generated by each store and city?

Store City	Store Id	₹
Abu Dhabi	A2001	3,612,980
	A2002	2,302,990
	A2005	1,691,185
	A2003	988,516
	A2006	732,051
Al Ain	A2009	309,756
Doha	A2023	3,094,129
Dubai	A2011	4,566,642
	A2010	3,630,100
	A2038	2,188,619
	A2013	1,639,011
	A2035	918,589
Fujairah	A2020	2,382,103
other	A2014	1,601,651
	A9002	701,988
	A9003	304,058
	A9001	27,611
RAK	A2026	1,911,228
Sharjah	A2027	2,719,571
	A2029	2,663,437

Fig represents Total revenue wrt Store id & Store city

**Insight:** Al Ain city has minimum revenue with only 1 store of 3 lakh\$. Rest all cities have more than 1 store having above 15 lakh \$. Dubai has 4 of 5 stores 16 lakh \$ - 45 lakh \$.

Which product categories contribute the most to total revenue?

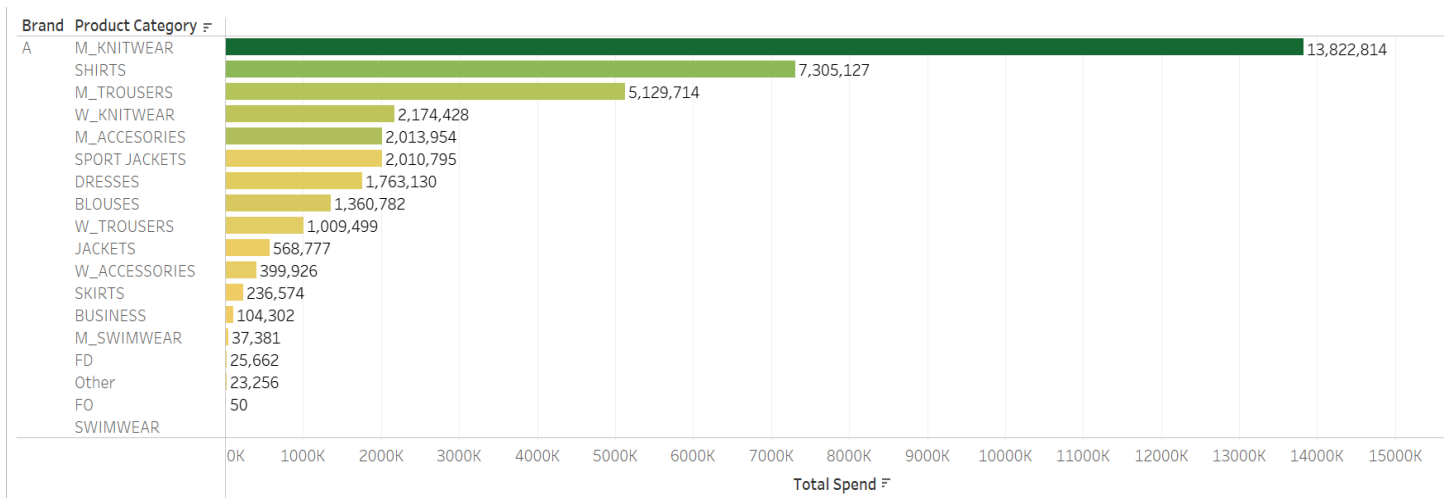


Fig tells revenue contribution by different product category

**Insight:** M\_Knitwear contributed the highest revenue of 1.38cr. Shirts, M\_trousers had 73 lakh , 51 lakh respectively. W\_knitwear, M\_Accesories & Sport jackets each make 0.2cr.Next 3 categories make 0.1cr each

What is the revenue trend over time?

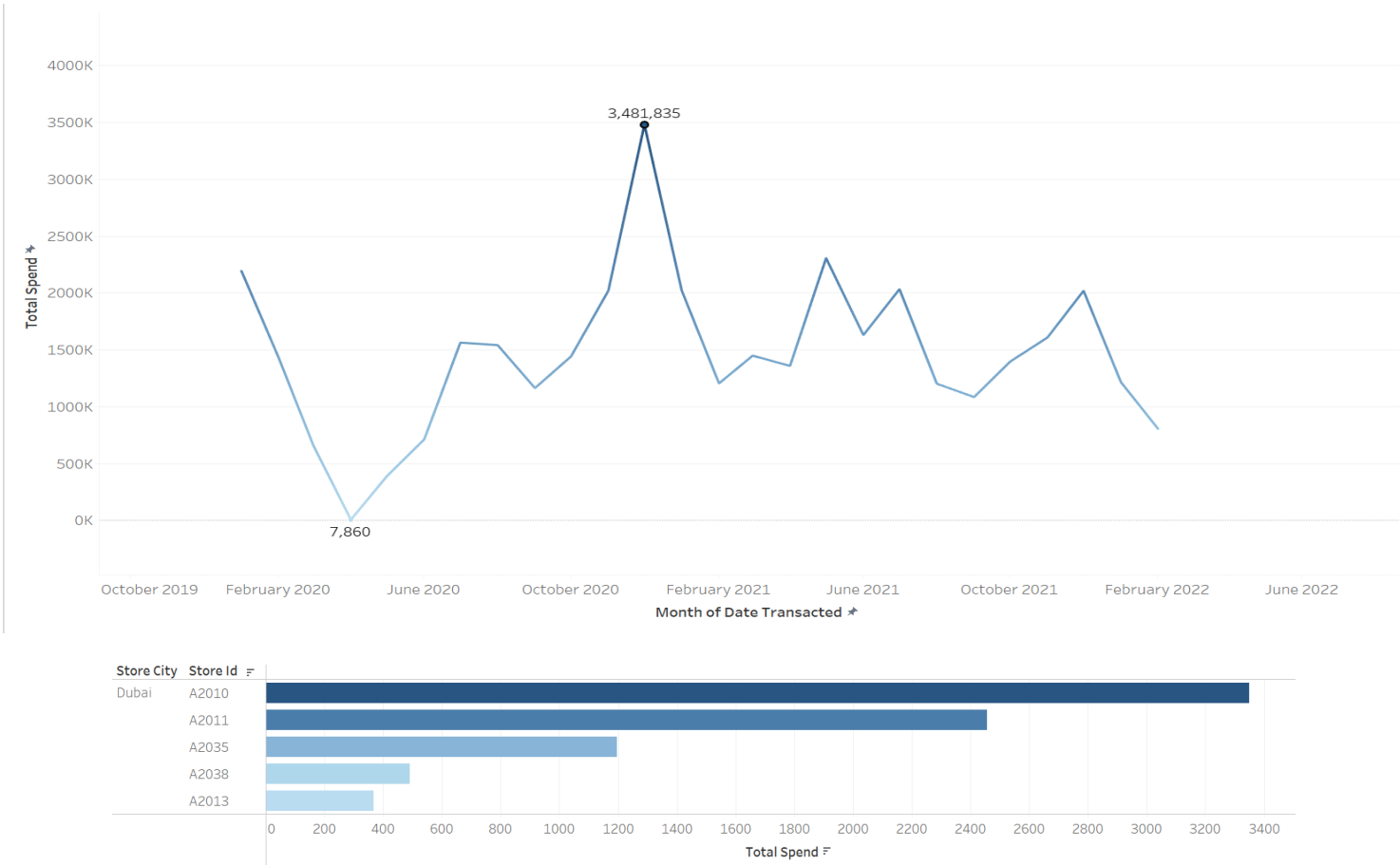


Fig shows sales pattern in april 2020

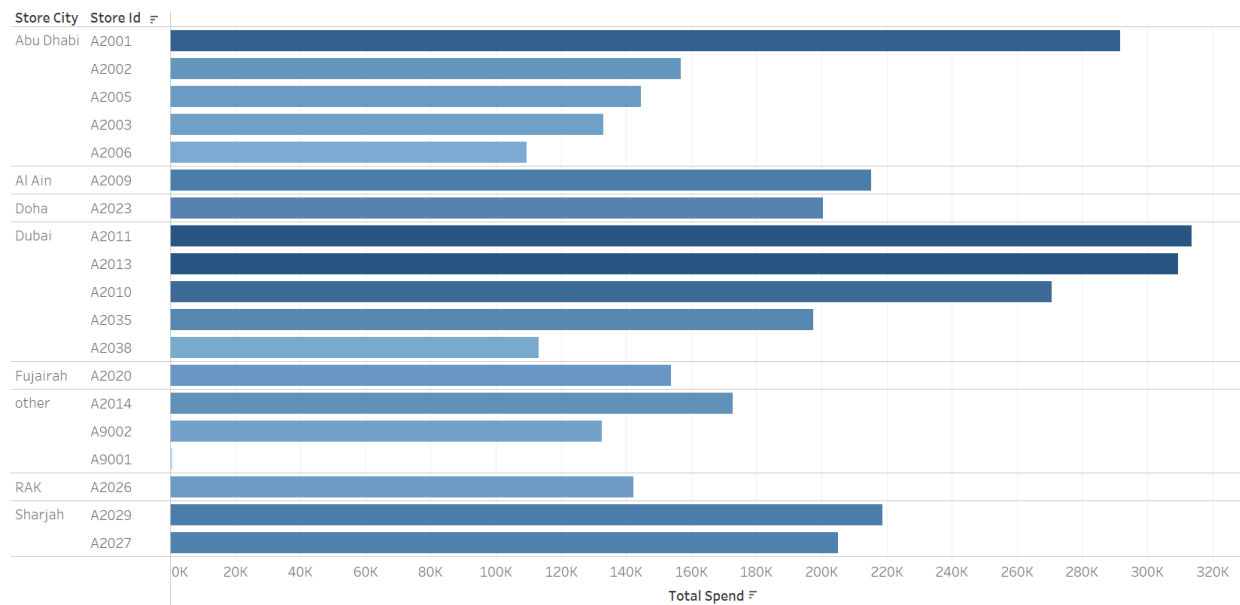


Fig shows sales pattern in December 2020

**Insight:** Started in 2020 with 2000k sales, in April 2020 had had lowest signup of 7 & sales dropped to 7k. As sales occurred only in Dubai stores only. Again December 2020 had an all time high revenue of 0.34cr. During this time all stores in all cities were doing Sales above 140k. Later on months sales stabilized to 1500k to 2000k.

## 2. Product Performance:

Which items have the highest sales volume and total spend?

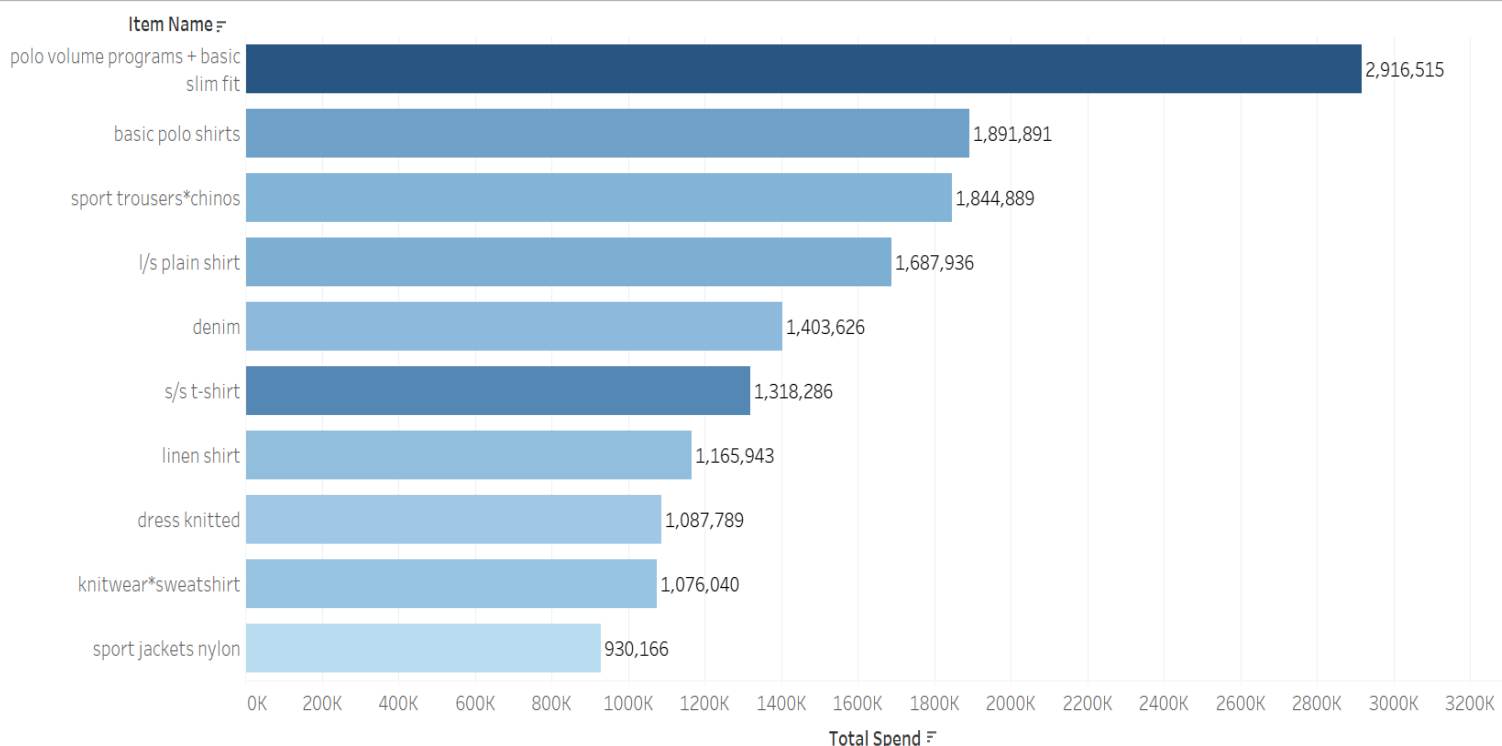
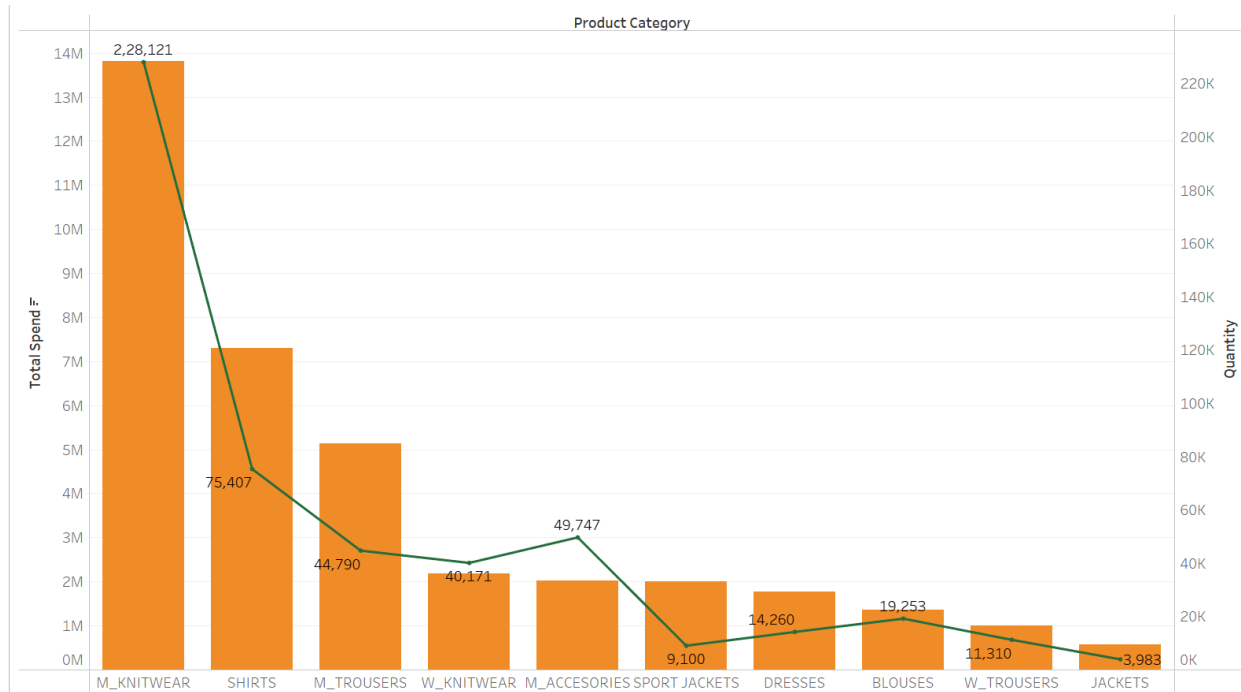


Fig shows total spend & color indicates sales volume per product item.

**Insight:** Products like polo volume programs slim fit alone contribute 29 lakhs, basic polo shirts & sport trousers send 18 lakhs. Rest products from top 10 make avg revenue above 10 lakhs.

How do different product categories perform in terms of quantity sold and revenue generated?



Green: Quantity, Orange: Total spend

**Insight:** M\_Knitwear contributed the highest revenue of 1.38cr. With quantity of 2.28lakh pcs, Shirts, M\_trousers had 73 lakh , 51lakh sales respectively but quantity is 75k & 44k. W\_knitwear, M\_Accesories & Sport jackets each make 0.2cr.and volume is 40k, 50k, 9k resp. Next 3 categories had lower volumes of 14k, 19k & 11k respectively.

What are the top-selling items in each store or city?



Store City	Store Id	
Abu Dhabi	A2001	women sneaker
	A2002	women sneaker
	A2003	watch
	A2005	watch
	A2006	watch
Al Ain	A2009	tank tops
Doha	A2023	women sneaker
Dubai	A2010	women sneaker
	A2011	women sneaker
	A2013	watch
	A2035	tank tops
	A2038	watch
Fujairah	A2020	watch
other	A2014	watch
	A9001	tank tops
	A9002	watch
	A9003	watch
RAK	A2026	women sneaker
Sharjah	A2027	watch
	A2029	watch

Fig is of top selling items at each store wrt city

**Insight:** Most stores have Watch, women sneaker and Tank tops as top-selling items

## 2. Sales Trends and Patterns:

What are the peak sales periods (e.g., days, months)?

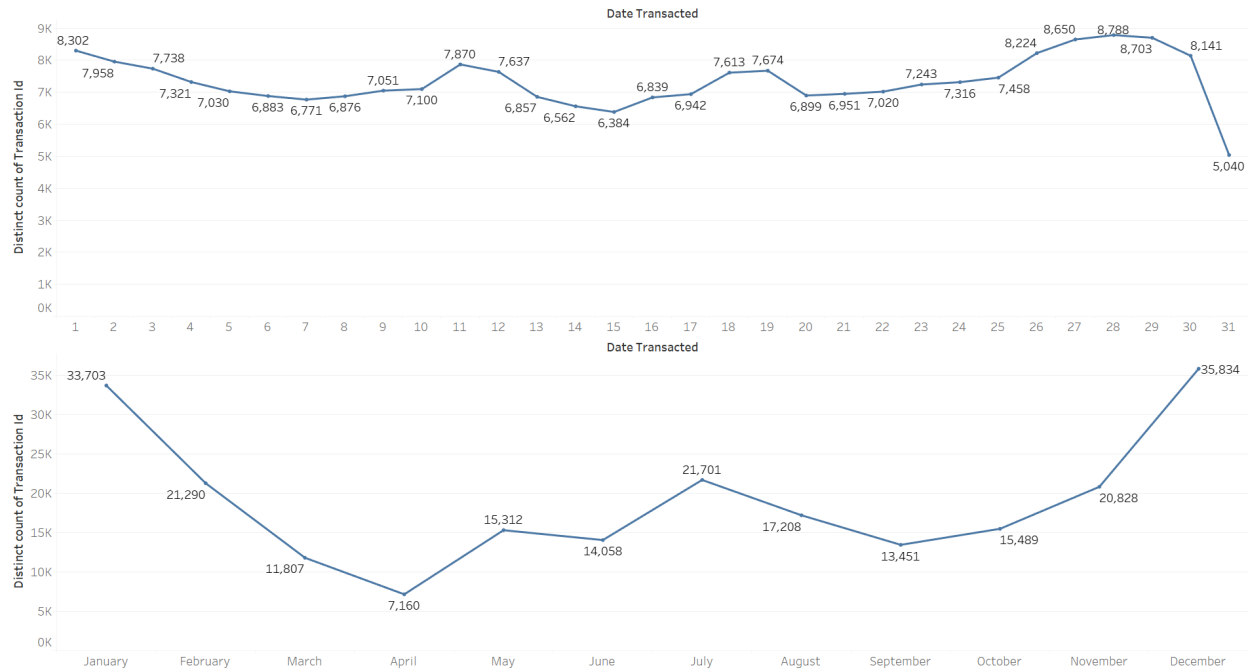
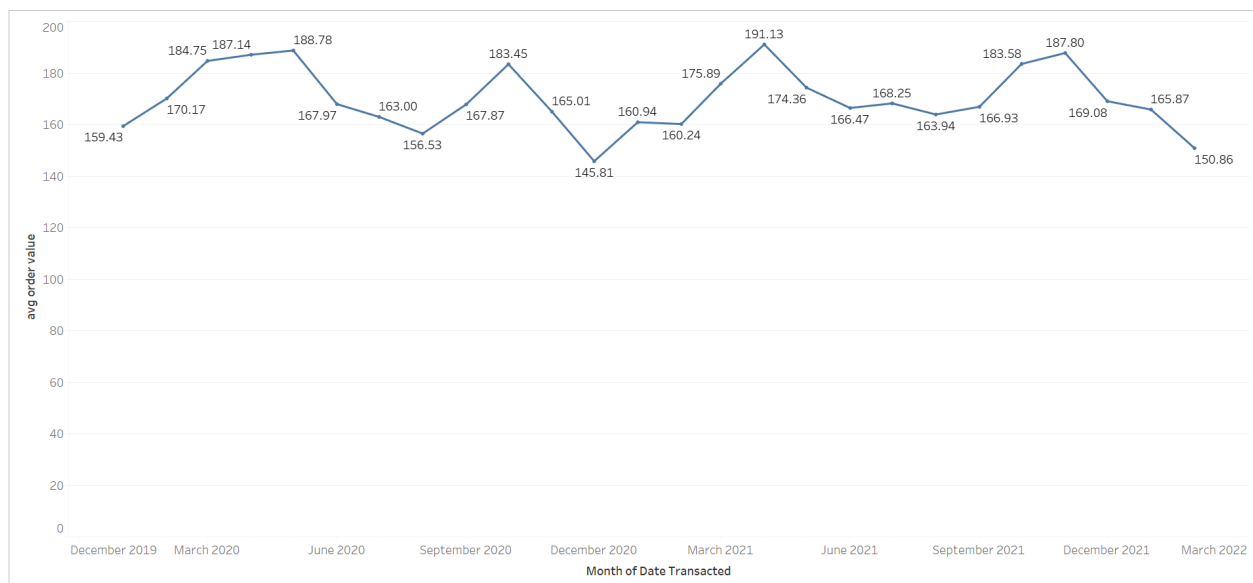


Fig tells the trend of sales volume in terms of days & months.

**Insight:** In terms of days only 31st have least volume of 5k transactions but rest all days avg of 7k-8k is maintained. In terms of months Dec & Jan have the highest transaction of 30k. But March & April had min of 11k & 7k transactions.

How does the average transaction value change over time?



**Insight:** During May 2020, April 2021, Nov 2021 with highest AOV of 188,191& 187 respectively. Lowest AOV was seen in months of Dec 2020 & March 2022 with AOV of 145 & 150 resp.

What is the correlation between signed up app date and total spend?

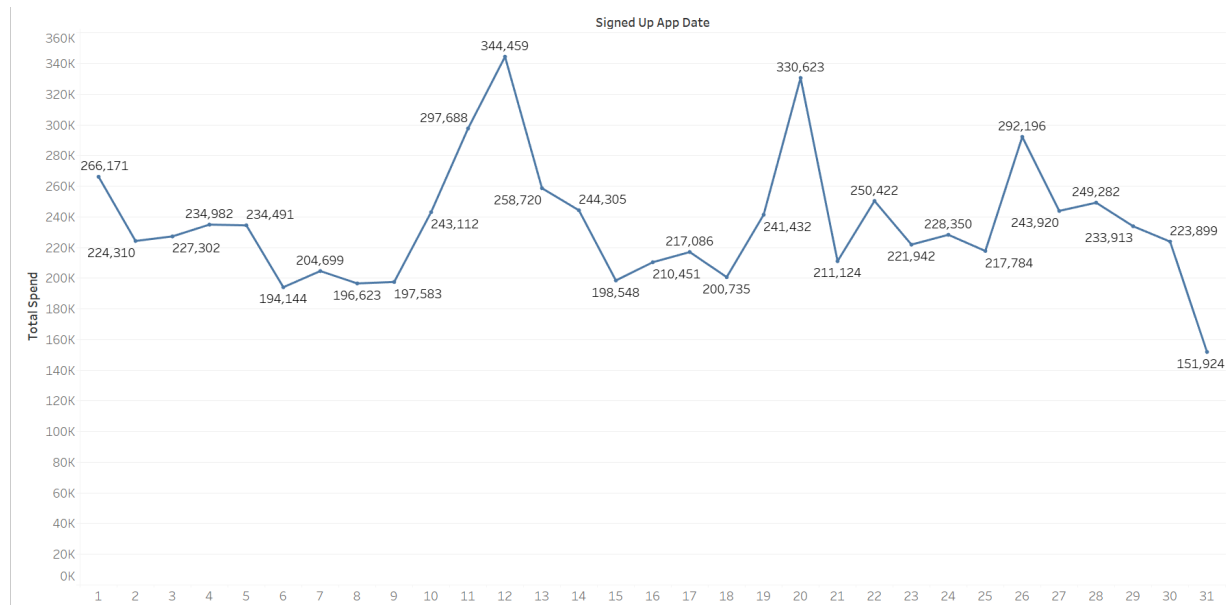
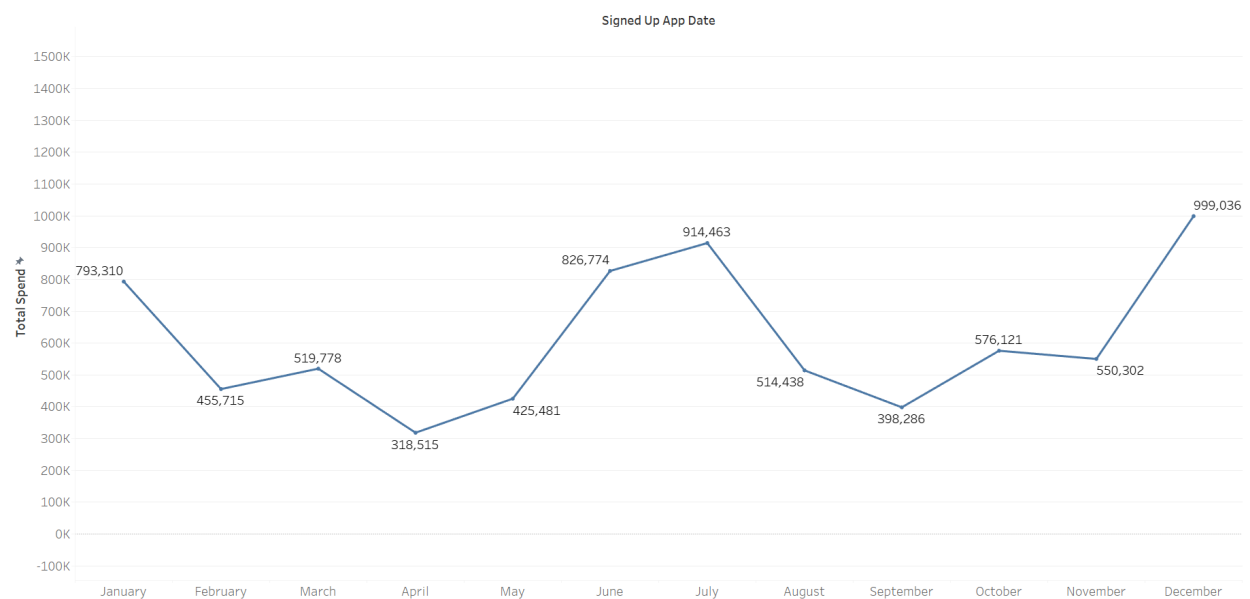


Fig shows Trend line of total spend & Signed up date.



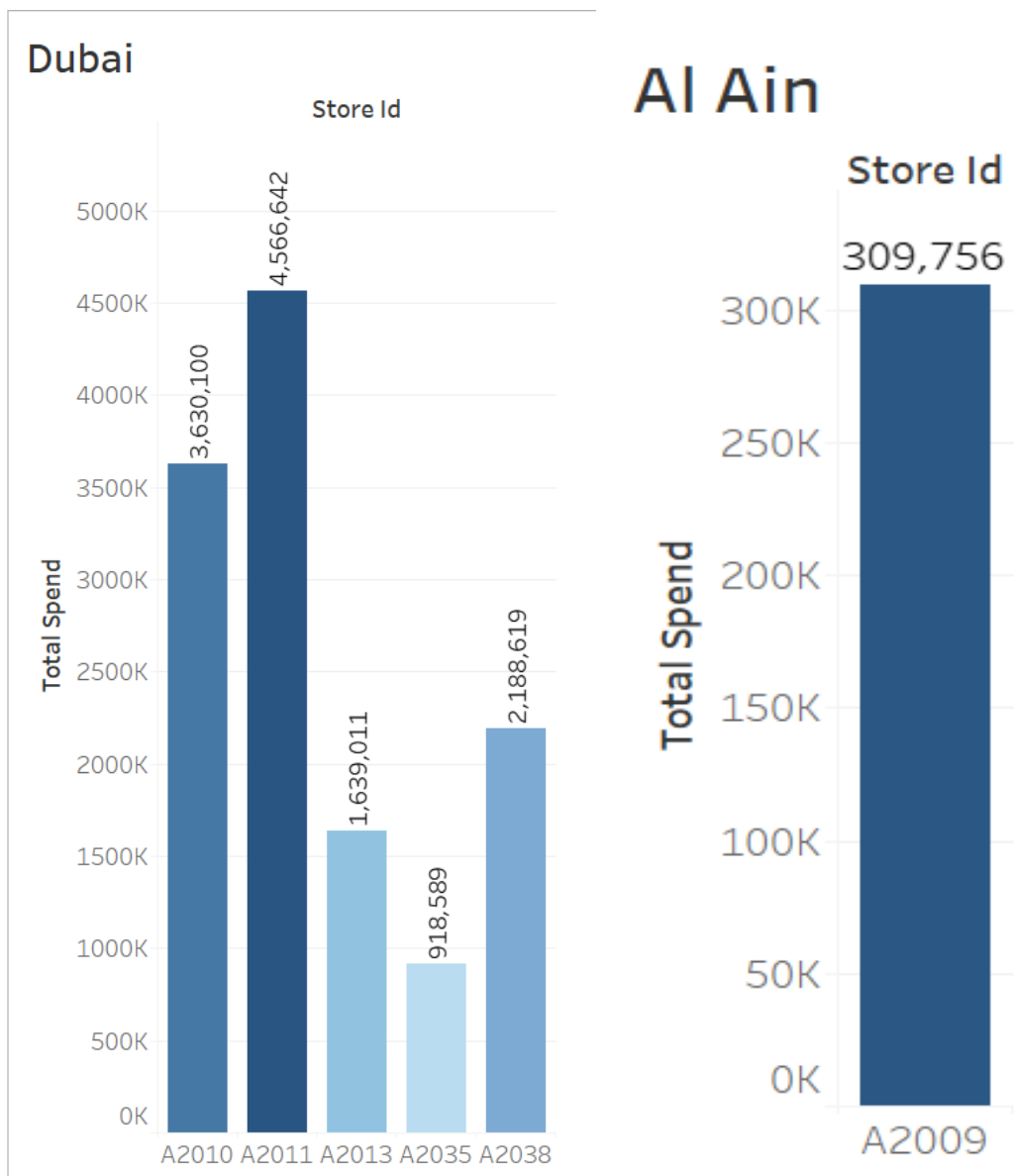
Relationship between signup month and total spend

**Insight:** As no. of transactions are least on 31st, so, Total spend is also 1.5 lakh. Similarly on 12th & 20th had peak revenue of 3.44 & 3.3lakh resp. In terms of months, April had the least revenue of 3.18 lakh. But in July & dec have highest sales of 9 lakh & 1cr.

Operational Insights

1. Store Performance:

Which stores have the highest sales and revenue?



**Insight:** A2011 store in Dubai has the highest revenue of 45 lakh. And Al ain has a single store but that is also performing below average making 3 lakhs.

How does store performance vary by city ?

Store City	Store Id	₹	Store City	Store Id	₹
Abu Dhabi	A2001	3,612,980	Abu Dhabi	A2001	19,172
	A2002	2,302,990		A2002	12,396
	A2005	1,691,185		A2005	8,934
	A2003	988,516		A2003	7,134
	A2006	732,051		A2006	5,379
Al Ain	A2009	309,756	Al Ain	A2009	2,572
Doha	A2023	3,094,129	Doha	A2023	17,429
Dubai	A2011	4,566,642	Dubai	A2011	24,211
	A2010	3,630,100		A2010	22,762
	A2038	2,188,619		A2038	16,675
	A2013	1,639,011		A2013	8,848
	A2035	918,589		A2035	6,366
Fujairah	A2020	2,382,103	Fujairah	A2020	10,536
other	A2014	1,601,651	other	A2014	13,581
	A9002	701,988		A9002	6,751
	A9003	304,058		A9003	3,313
	A9001	27,611		A9001	60
RAK	A2026	1,911,228	RAK	A2026	10,836
Sharjah	A2027	2,719,571	Sharjah	A2027	15,931
	A2029	2,663,437		A2029	14,955

fig tells revenue & other tells sales volume.

**Insight:** Stores in Dubai have higher customers likewise their transaction frequency & revenue are also higher. Similar trend is followed in Abu Dhabi stores. Sharjah has 2 stores with 6k & 7k customers contributing revenue of 26 lakh & 27 lakh with 14k transaction & 16k transaction.

## Comparative Analysis

### 1. Cross-Store Comparisons:

How do sales and customer demographics differ between stores in different cities?

Store City	Store Id	Revenue	Store City	Store Id	Revenue
Abu Dhabi	A2001	8,093	Abu Dhabi	A2001	3,612,980
	A2002	4,871		A2002	2,302,990
	A2005	3,929		A2005	1,691,185
	A2003	3,208		A2003	988,516
	A2006	2,722		A2006	732,051
Al Ain	A2009	484	Al Ain	A2009	309,756
Doha	A2023	7,486	Doha	A2023	3,094,129
Dubai	A2011	8,426	Dubai	A2011	4,566,642
	A2010	6,856		A2010	3,630,100
	A2038	3,009		A2038	2,188,619
	A2013	1,045		A2013	1,639,011
	A2035	635		A2035	918,589
Fujairah	A2020	3,958	Fujairah	A2020	2,382,103
other	A9002	37	other	A2014	1,601,651
	A9003	11		A9002	701,988
	A2014	2		A9003	304,058
	A9001	0		A9001	27,611
RAK	A2026	2,834	RAK	A2026	1,911,228
Sharjah	A2029	7,569	Sharjah	A2027	2,719,571
	A2027	6,287		A2029	2,663,437

Fig represents Total customers & revenue wrt Store id & Store city

**Insight:** Al ain city with least number of customers of 484 doing 3 lakh revenue. Highest revenue was made by A2011 having 8.4k customers generating 45 lakh revenue. Similarly, the second highest A2001 store in Abu Dhabi is making revenue of 36 lakh from 8k customers. This tells that Number of customers is directly proportional to revenue generated.

What are the key performance differences between high-performing and low-performing stores?

Store City	Store Id	
Dubai	A2011	Total spend: 4,566,642 Sales Volume: 24,211
other	A9001	Total spend: 27,611 Sales Volume: 60

**Insight:** Highest performing store is A2011 in Dubai with total revenue of 45 lakhs with transaction frequency of 24k. Whereas, store A9001 is lowest performing with total revenue of 27k from only 60 transactions.

## 2. Product Comparisons:

What are the differences in sales performance between high-level product categories?

High Level Product Category	₹
Shirts	7,330,789
Polo	6,462,740
Knitwear	6,295,702
Trousers	2,741,201
Tricot	2,242,950
Jackets	2,016,200
Accessories	1,908,201
Dresses	1,777,536
Denim	1,586,882
Blouses	1,360,782

**Insight:** Among high level product categories, Top 3 categories are driving sales. Shirts are leading with 73 lakhs, followed by Polo & knitwear with 64 & 62 lakh sales. Whereas, Trousers, Tricot & Jackets with 20 lakhs. Similarly accessories, dresses, denim, blouses make above 10 lakh sales.

## Recommendations:

### Customer Demographics and Sales Volume

1. **Targeted Marketing Campaigns:** Focus marketing efforts on the top three countries (UAE, India, Philippines) to boost sales further. Create country-specific promotions and advertisements.
2. **Loyalty Programs:** Promoting loyalty programs in these regions to retain existing customers and attract new ones. Has worked out in previous years also it will boost sales in further days as well.

### Customer Segments

3. **Upgrade Incentives:** Encourage Silver segment customers to move up to Gold or Platinum by offering special incentives for higher spending.
4. **Personalized Offers:** Provide personalized offers to Platinum customers to maintain their loyalty and encourage even higher spending.

### Average Order Value (AOV)

5. **City-Specific Strategies:** Investigate why Fujairah has the highest AOV and replicate successful strategies in other cities.
6. **Boost Al Ain Sales:** Introduce promotions or bundles in Al Ain to increase the AOV.

### Loyalty Program Sign Ups

7. **Consistent Engagement:** Maintain high engagement levels to keep the signup rate above 1500/month.
8. **Analyze Fluctuations:** Understand the reasons behind fluctuations and low signup periods to avoid similar occurrences in the future.



## Customer Lifetime Value (CLV)

9. **Focus on High CLV Regions:** Enhance customer service and engagement in regions with high CLV to maximize long-term value.
10. **Expand Successful Strategies:** Apply successful strategies from Cambodia to other regions.

## Store Performance

11. **Analyze Top Performers:** Study the operations of the highest-performing stores to identify best practices.
12. **Improve Low Performers:** Implement targeted improvements in the lowest-performing stores, such as staff training or localized marketing efforts.
13. **Optimize Al Ain Store:** Increase promotional activities and improve the product mix in the Al Ain store to boost its performance.

## Product Category Performance

14. **Promote Top Categories:** Increase marketing for top-performing categories to drive even higher sales.
15. **Diversify Product Range:** Introduce new products in the less performing categories to capture more market share.

Tools Used :



## Tableau Data Visualization: Analysis Links

- [Customer Demographics and Sales Volume](#)
- [Sales and revenue analysis](#)

Medium article link: [Apparel sales analysis](#)