

Project Description: This analysis on Amazon data from India was done using MS Excel and involved the following skills:

Data Cleaning
Data Manipulation and calculations
Data Transformation
Creating Pivot Tables and Charts
VLOOKUP and Relationships
Insights Analysis

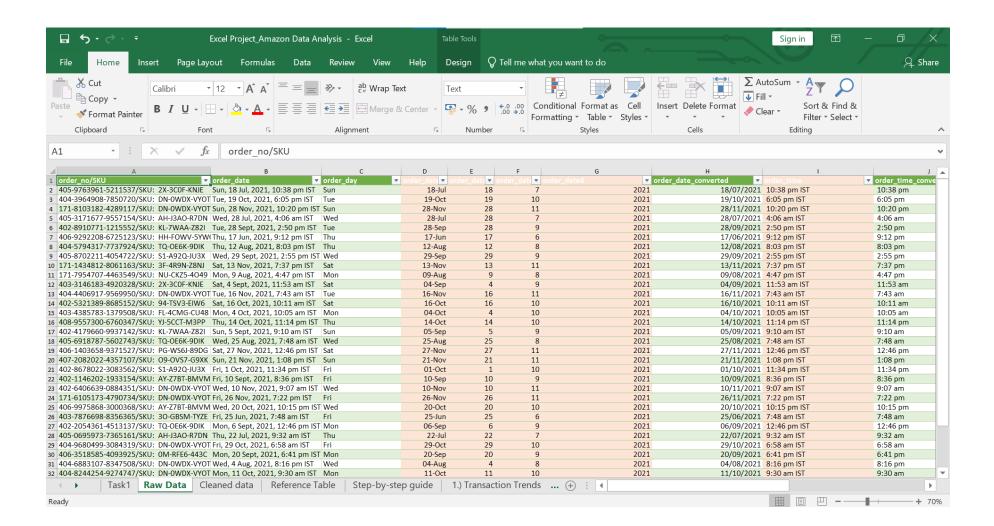
The following are the guide questions we want to answer for this project

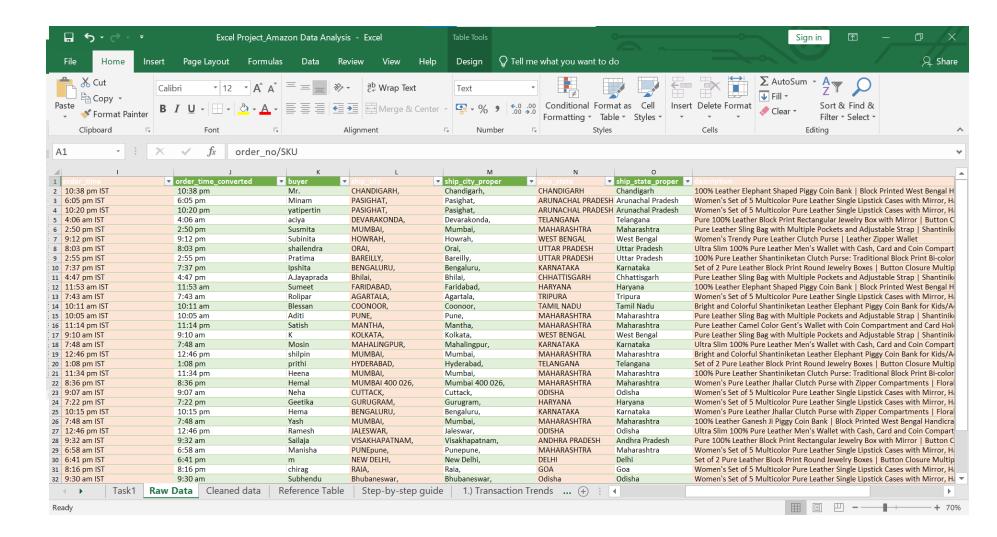
- A. Do we have a stable transaction trend?
- B. What is the distribution of the order status?
- C. Can we say something about the shipping fee? Do we think the shipping fee affects the number of orders?

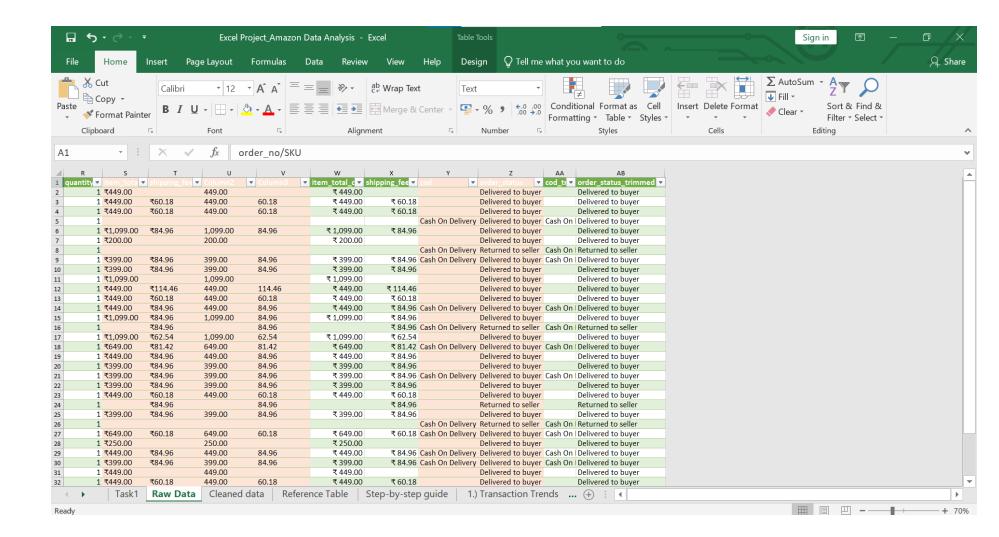
Raw Data

The images below show the raw data that I worked with. I has 28 columns and 172 rows. There are different highlights because I performed data cleaning and transformations. The cleaned data will be found in the next sheet called "Cleaned data" where all the pertinent information will only be included. The column descriptions are the following:

- A. order no/SKU: order ID
- B. order data: date and time the order was done
- C. order_day: extracted day from order_data
- D. order_data1: month and day only
- E. order_data2: only the day
- F. order_data3: transformed month name to number
- G. order_data4: extracted only the year of order data
- H. order_date_converted: date in number format
- I. order_time: extracted only the time from order_data
- J. order_time_converted: extracted only the time without the timezone specifications
- K. buyer: customer name info
- L. ship_city: delivery destination (city)
- M. ship city proper: transformed ship city into proper capitalization
- N. ship_state: delivery destination (state)
- O. ship_state_proper: transformed ship_state into proper capitalization
- P. description: product description
- Q. description trimmed: trimmed spaces description
- R. quantity: product quantity
- S. item_total: item total price (raw text)
- T. shipping_fee: shipping fee (raw text)
- U. Extracted only number
- V. Extracted only number
- W. item_total_cost: transformed total_cost to currency format
- X. shipping_fee_cost: transformed shipping_fee to currency format
- Y. cod: delivery type
- Z. order_status: delivered to buyer or returned to seller
- AA. cod trimmed
- AB. order_status trimmed

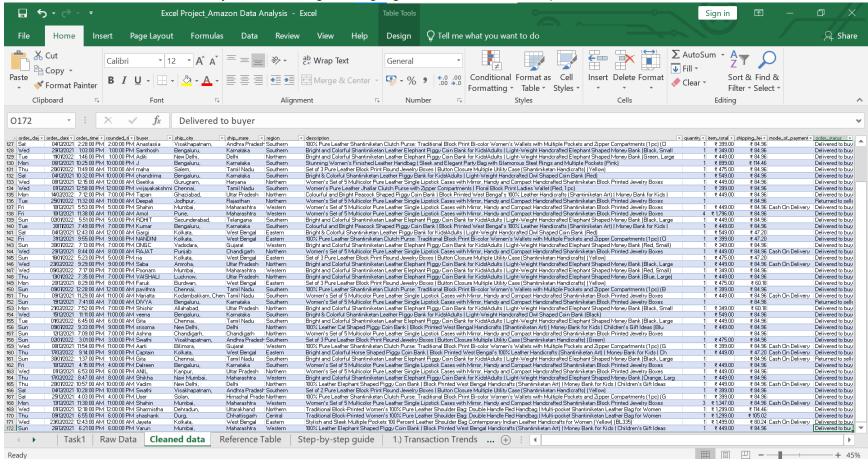






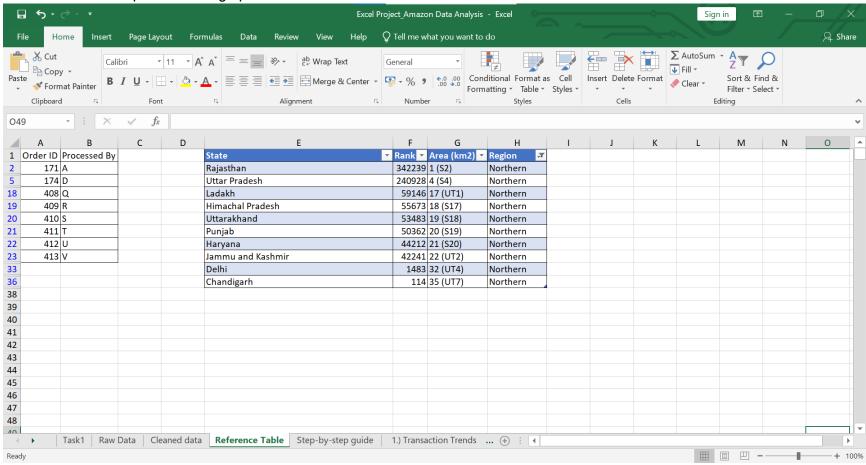
Cleaned data:

Below is the cleaned data that only included the green-highlighted columns from the previous sheet.



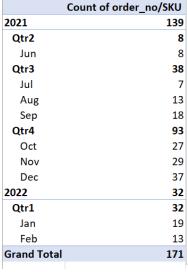
Reference Table

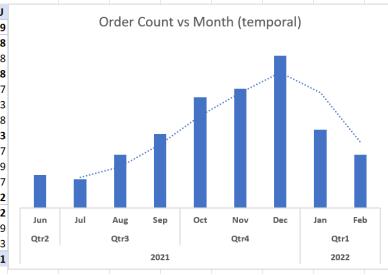
These tables will help us in looking up information about certain data we have.



ANALYSIS PROPER

Transaction Trends





Title

Order count by month, Jun 2021-Feb 2022 (temporal)

Description

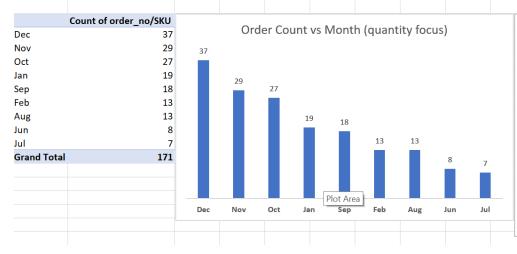
Here is a simple overview of order count by month. We will analyse further in the following tables and charts.

Insigh

There is an obvious imbalance of order count among months in the given period.

Assumption

There are certain months and quarters that reflect the greatest number of orders.



Title

Order count by month, Jun 2021-Feb 2022 (quantitative)

Description

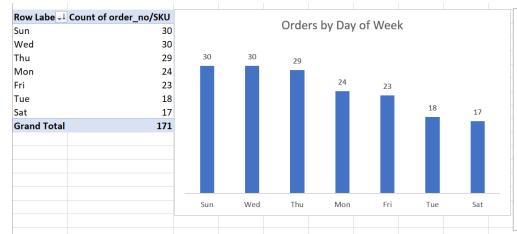
For the entire scope of the data, December was the month that got the most orders with 37 counts. July, on the other hand, had the least with just 7 orders. This was counted irresspective of quarters and years.

Insight

Seeing the top 3 months in terms of order counts, they also fall down in Q4 of 2021 as reflected in the above chart.

Assumption

Q4 of the year tends to garner the most orders. This could reflect the purpose of buying items by Indians. It is wrong to assume Christmas season gift buying because only a minority of Indians celebrate the said holiday. We will see what purpose/s they have based on the items they bought.



Title

Order count by Day of Week

Description

Sundays and We Plot Area have the most accumulated orders among other days of the week with 30 orders each for the entire period

Insight

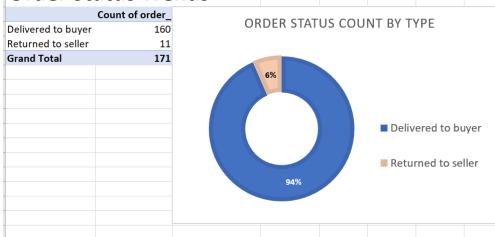
Sundays and Wednesdays are the most active days that customers check-out with their orders

Assumption/Recommendation

The company might want to create reminders through app or email that will prompt pitches to customers about good deals and/or checking out their 'added to cart' items on the lesser active days (e.g., Saturdays and Tuesdays).



Order Status Trends



Title

Order Status Count by Type

Description

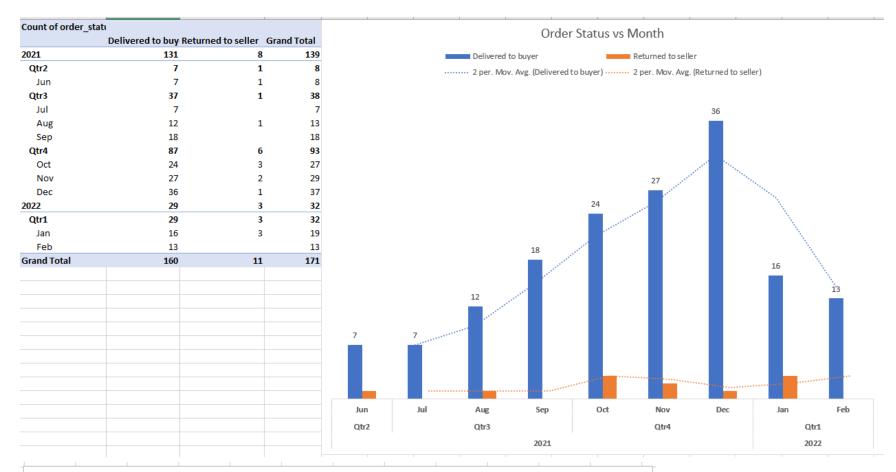
94% or 160 out of 171 orders have concluded where the package have been delivered to the buyer. While only 6% or 11 orders have been returned to the seller.

Insight

We can say that the delivery system has great efficiency at 94%. Still, we want to understand why certain orders have been returned to the seller and optimize our delivery services.

Assumption

We will understand this trend by looking into other elements involving returned orders.



Title

Order Status by Month

Description

Throughout all months, a huge majority of orders have been delivered to the buyer as signified by the blue line running over the orange line signifying the trend for orders being returned to the seller.

Insight and Assumption

Again, we see that the system is efficient in delivering orders. But we want to improve this as well to possibly 100% delivered to buyer status. Maybe it is because of the product type. We shall see this in the next chart.

Insight: After manual comparison, there is no unique item that was returned to the seller that is not delivered to the buyer. We shall see if another variable, the region, plays a role in returning orders to sellers.

	Count of region	Insight:													1
Delivered to buyer	160		s returned to	the seller	were from	the North	ern region	with 4 cour	ts, followe	d by South	ern region	with 3 cou	nts. This va	riable	
Central	7	seems to provide					0		•	,					
Eastern	22														
Northeastern	6	Assumption:													
Northern	36	Both South and North regions may be far and many factors could have played a role in the customer choice to return the order to the seller.								er.					
Southern	51	-increased deliver	y time												
Western	38					_									
Returned to seller	11	-exceeding expect	ted date of ar	rival hence	e irritating	the custom	er.								
Eastern	1														
Northeastern	1														
Northern	4														
Southern	3	regions so that:	سنف سمينا مامام												
Western	2	-there is decrease			chinning ti	mo									
Grand Total	171	-less damage to the packages due to less shipping time -orders arrive closer to the expected date of delivery.													
		orders dirive cio.	ser to the exp	corca dare	. Or deliver	γ.									
		Doing so may not only increase delivery rate and that we expect the order status of 'delivered to buyer' be higher than 94%, but									t we can als	so lower			
		the shipping fee in exchange to enticing more customers to avail our delivery services. The balance between these two variables shall be seen if ever													
		the recommendat	_	_				,							
I					-										_

Shipping Fee Analysis



Title

Average Shipping Fee by Region

Description

The Southern region has the highest average shipping fee of $\stackrel{?}{\sim}$ 93. While the lowerst is the Eastern region with $\stackrel{?}{\sim}$ 59.51.

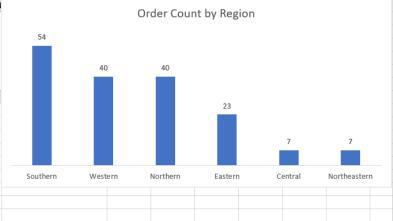
Insight

Areas nearer the Eastern region have lower shipping fees, where Southern and Northern shipping fees are the highest.

Assumption

The storehouse or base must be located in the Eastern region that is why the pricing is like this. But does this affect the willingness to order by the customers? We shall see in the next chart/s.

ĺ		Count of order_no/SKU
	Southern	54
	Western	40
	Northern	40
	Eastern	23
	Central	7
	Northeastern	7
	Grand Total	171



Title

Order Count by Region

Description

The Southern region has the highest total of orders with 54 orders throughout the period; while both Central and Northeastern regions have the least with 7 orders.

Insight

The Southern region still has the most orders, followed by a tie between the Western and Northern regions. While it is reasonable for people to buy more if the shipping fee is less, this does not appear to be the case since the Southern region having the highest shipping fee yet garnering the most orders.

Assumption

We'd like to see if there is a correlation between shipping fee and order count in our next analysis.

