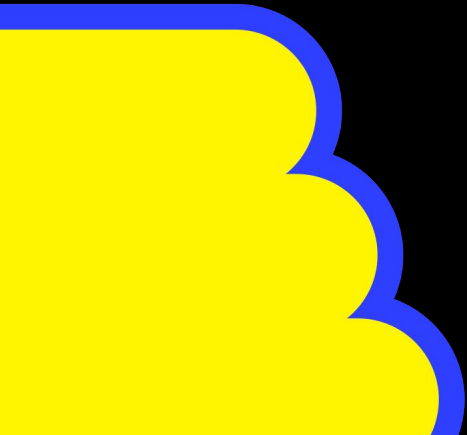
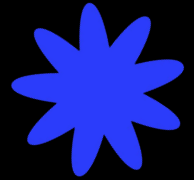
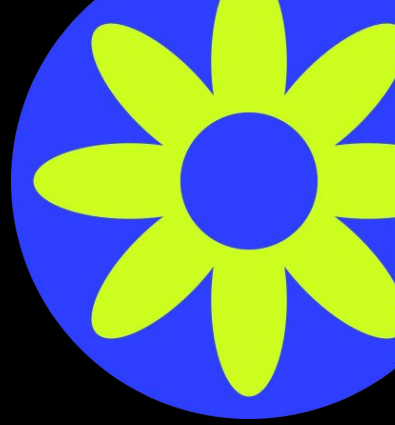


AMAZON * E-COMMERCE SALES ANALYSIS

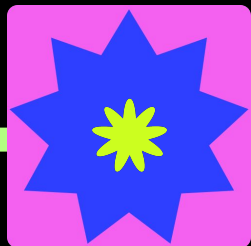




agenda

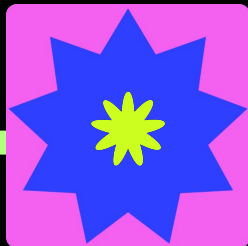


Introduction



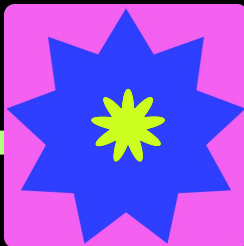
Introducing
presenters
Defining our data,
variables, etc

Sabrina: State + Service type



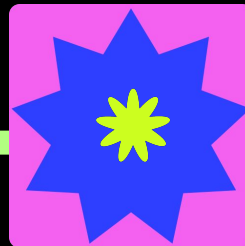
Compare how
profitability can be
affected by where a
seller ships to as
well as compare ship
service type

Ariel: Clothing type



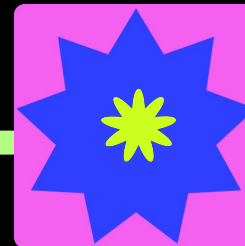
Compare the most
ordered clothing types,
and checking their
relationship to status
type

Brianna + Time Variable



Finding relationship
between average orders
and order statuses
throughout time

Robert M. + Successfully Delivered



Compare retailer
order outcomes,
promotion usage, and
how they affect
profitability

```

1 index,Order ID,Date,Status,Fulfilment,Sales Channel ,ship-service-level,Style,SKU,Category,Size,ASIN,Courier Status,Qty,currency,Amount,ship-city,ship-state,ship-f
2 0,405-8078784-5731545,04-30-22,Cancelled,Merchant,Amazon.in,Standard,SET389,SET389-KR-NP-S,Set,S,B09KXVBD7Z,,0,INR,647.62,MUMBAI,MAHARASHTRA,400081.0,IN,,False,Eas
3 1,171-9198151-1101146,04-30-22,Shipped - Delivered to Buyer,Merchant,Amazon.in,Standard,JNE3781,JNE3781-KR-XXXL,kurta,3XL,B09K3WFS32,Shipped,1,INR,406.0,BENGALURU,
4 2,404-0687676-7273146,04-30-22,Shipped,Amazon,Amazon.in,Expedited,JNE3371,JNE3371-KR-XL,kurta,XL,B07WV4JV4D,Shipped,1,INR,329.0,NAVI MUMBAI,MAHARASHTRA,410210.0,IN
5 3,403-9615377-8133951,04-30-22,Cancelled,Merchant,Amazon.in,Standard,J0341,J0341-DR-L,Western Dress,L,B099NRCT7B,,0,INR,753.33,PUDUCHERRY,PUDUCHERRY,605008.0,IN,,F
6 4,407-1069790-7240320,04-30-22,Shipped,Amazon,Amazon.in,Expedited,JNE3671,JNE3671-TU-XXXL,Top,3XL,B098714BZP,Shipped,1,INR,574.0,CHENNAI,TAMIL NADU,600073.0,IN,,Fa
7 5,404-1490984-4578765,04-30-22,Shipped,Amazon,Amazon.in,Expedited,SET264,SET264-KR-NP-XL,Set,XL,B08YN7XDSG,Shipped,1,INR,824.0,GHAZIABAD,UTTAR PRADESH,201102.0,IN,
8 6,408-5748499-6859555,04-30-22,Shipped,Amazon,Amazon.in,Expedited,J0095,J0095-SET-L,Set,L,B08CMHNW8N,Shipped,1,INR,653.0,CHANDIGARH,CHANDIGARH,160036.0,IN,IN Core
9 7,406-7807733-3785945,04-30-22,Shipped - Delivered to Buyer,Merchant,Amazon.in,Standard,JNE3405,JNE3405-KR-S,kurta,S,B081WX4G4Q,Shipped,1,INR,399.0,HYDERABAD,TELAN
10 8,407-5443024-5233168,04-30-22,Cancelled,Amazon,Amazon.in,Expedited,SET200,SET200-KR-NP-A-XXXL,Set,3XL,B08L91ZZXN,Cancelled,0,,HYDERABAD,TELANGANA,500008.0,IN,IN
11 9,402-4393761-0311520,04-30-22,Shipped,Amazon,Amazon.in,Expedited,JNE3461,JNE3461-KR-XXL,kurta,XXL,B08B3XF5MH,Shipped,1,INR,363.0,Chennai,TAMIL NADU,600041.0,IN,,F
12 10,407-5633625-6970741,04-30-22,Shipped,Amazon,Amazon.in,Expedited,JNE3160,JNE3160-KR-G-S,kurta,S,B07K3YQLF1,Shipped,1,INR,685.0,CHENNAI,TAMIL NADU,600073.0,IN,,Fa
13 11,171-4638481-6326716,04-30-22,Shipped,Amazon,Amazon.in,Expedited,JNE3500,JNE3500-KR-XS,kurta,XS,B098117DJ3,Shipped,1,INR,364.0,NOIDA,UTTAR PRADESH,201303.0,IN,,F
14 12,405-5513694-8146768,04-30-22,Shipped - Delivered to Buyer,Merchant,Amazon.in,Standard,JNE3405,JNE3405-KR-XS,kurta,XS,B081XCMYXJ,Shipped,1,INR,399.0,Amravati,MA
15 13,408-7955685-3083534,04-30-22,Shipped,Amazon,Amazon.in,Expedited,SET182,SET182-KR-DH-XS,Set,XS,B085HS947T,Shipped,1,INR,657.0,MUMBAI,MAHARASHTRA,400053.0,IN,,Fal
16 14,408-1298370-1920302,04-30-22,Shipped - Delivered to Buyer,Merchant,Amazon.in,Standard,J0351,J0351-SET-L,Set,L,B09CSSQY4F,Shipped,1,INR,771.0,MUMBAI,MAHARASHTRA,
17 15,403-4965581-9520319,04-30-22,Shipped - Delivered to Buyer,Merchant,Amazon.in,Standard,PJNE3368,PJNE3368-KR-6XL,kurta,6XL,B09PY99SVJ,Shipped,1,INR,544.0,GUNTAKAL
18 16,406-9379318-6555504,04-30-22,Shipped,Amazon,Amazon.in,Expedited,JNE3721,JNE3721-KR-XXL,kurta,XXL,B099FCT65D,Shipped,1,INR,329.0,JAIPUR,RAJASTHAN,302020.0,IN,IN
19 17,405-9013803-8009918,04-30-22,Shipped,Amazon,Amazon.in,Expedited,JNE3405,JNE3405-KR-XL,kurta,XL,B081WT6GG7,Shipped,1,INR,399.0,NEW DELHI,DELHI,110074.0,IN,,False
20 18,402-4030358-5835511,04-30-22,Shipped - Delivered to Buyer,Merchant,Amazon.in,Standard,JNE3697,JNE3697-KR-XXL,kurta,XXL,B098133PV5,Shipped,1,INR,458.0,Gurgaon,HA
21 19,405-5957858-1051546,04-30-22,Shipped,Amazon,Amazon.in,Expedited,SET254,SET254-KR-NP-XS,Set,XS,B0983DDPL6,Shipped,1,INR,886.0,BENGALURU,KARNATAKA,560017.0,IN,,Fa
22 20,405-0607769-0716360,04-30-22,Shipped,Amazon,Amazon.in,Expedited,JNE3795,JNE3795-KR-XXXL,kurta,3XL,B09HMXJVFS,Shipped,1,INR,517.0,TIRUCHIRAPPALLI,TAMIL NADU,6200
23 21,404-8494550-5860325,04-30-22,Shipped,Amazon,Amazon.in,Expedited,SET345,SET345-KR-NP-M,Set,M,B09KXV4BN8,Shipped,1,INR,666.0,BENGALURU,KARNATAKA,560040.0,IN,IN Cc

```

Our data:

E-Commerce Sales Dataset

What is the data about?

What can we analyze?

What are some variables we can use?

SOURCE:

Cleaning the data

What were some variables we were focusing on?

What work did we do to clean the data?



Cleaning: Dropping Irrelevant Dates

```
# find max and min dates
print(f"Min Date: {consumer_sale_df['Date'].min()}")
print(f"Max Date: {consumer_sale_df['Date'].max()}")
```

Min Date: 2022-03-31 00:00:00

Max Date: 2022-06-29 00:00:00

```
# drop March values (just last day of March (outlier) with only few rows of data in March)
consumer_sale_df = consumer_sale_df[consumer_sale_df['Date'] != "2022-03-31"]

# updated min and max dates
print(f"Min Date: {consumer_sale_df['Date'].min()}")
print(f"Max Date: {consumer_sale_df['Date'].max()}")
```

Min Date: 2022-04-01 00:00:00

Max Date: 2022-06-29 00:00:00



**How do average total
orders and order
statuses vary across
weekdays and months
within the quarter?**



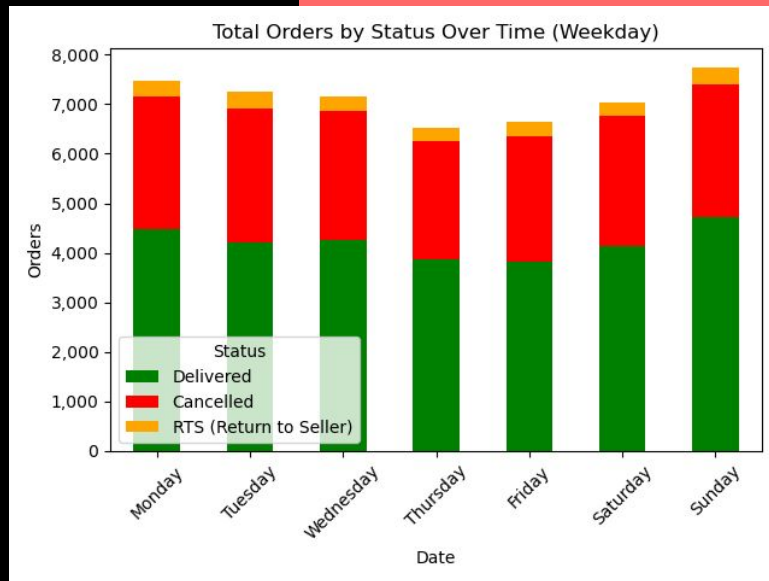
Weekday: Total Orders

Difference in **average total orders** across different weekdays? ❌


- F-Statistic: 0.44
- P-value: 0.85

High p-value (> 0.05) and low F-stat indicate that there is no statistically significant difference across weekdays.

Day	
Monday	7477
Tuesday	7247
Wednesday	7165
Thursday	6513
Friday	6634
Saturday	7040
Sunday	7735

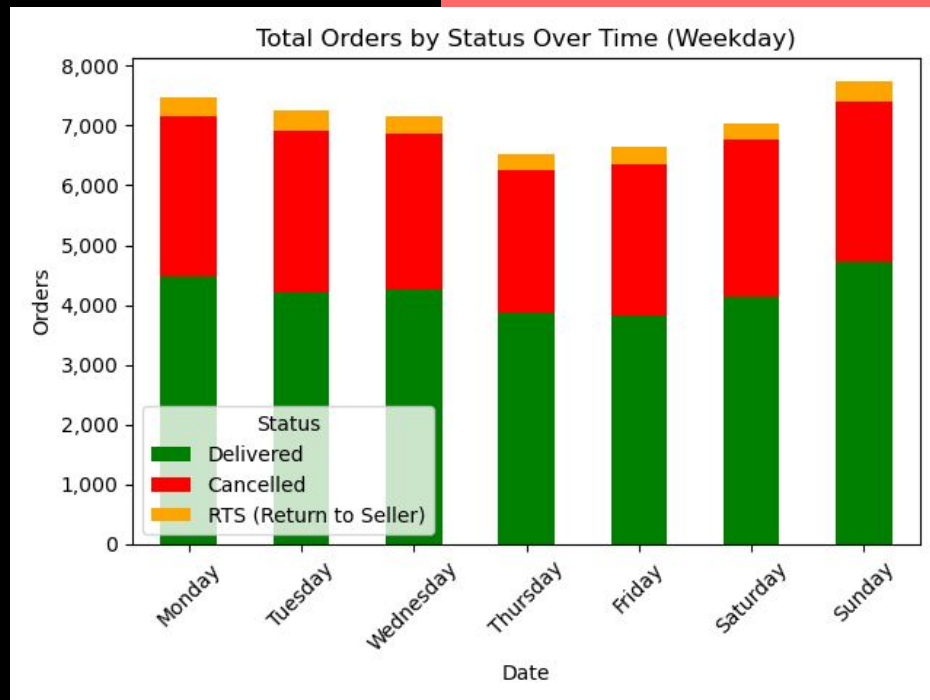


Weekday: Order Status

Relationship between
order status and
weekdays? 

- Chi-square statistic: 33.12
- P-value: $9.27e-4$

Low p-value (< 0.05) indicates statistically significant relationship between order status (delivered, cancelled, or returned) and weekday.



Weekday: Order Status

Cancellations:

- Worst: **Fridays** (↑ 4.6%)
- Best: **Sundays** (↓ 5%)

Successful Deliveries:

- Worst: **Fridays** (↓ 2.7%)
- Best: **Sundays** (↑ 3%)


Returned:

- Worst: **Tuesdays** (↑ 10%)
- Best: **Saturdays** (↓ 10.7%)

Status	Total Orders	Success Rate (%)	Cancelled Rate (%)	Returned Rate (%)
Day				
Monday	7477	59.957202	35.789755	4.253043
Tuesday	7247	57.982613	37.380985	4.636401
Wednesday	7165	59.399860	36.273552	4.326588
Thursday	6513	59.235375	36.634423	4.130201
Friday	6634	57.552005	38.317757	4.130238
Saturday	7040	58.792614	37.443182	3.764205
Sunday	7735	60.969619	34.802844	4.227537

ORDER COUNT: OBSERVED - EXPECTED			
Status	Cancelled	Delivered	RTS (Return to Seller)
Day			
Monday	-61.959085	59.034531	2.924555
Tuesday	55.263275	-85.879866	30.616591
Wednesday	-24.709622	16.637610	8.072012
Thursday	1.042461	4.410712	-5.453173
Friday	112.734175	-107.182149	-5.552027
Saturday	58.063400	-26.402823	-31.660577
Sunday	-140.434603	139.381984	1.052619

Monthly: Total Orders

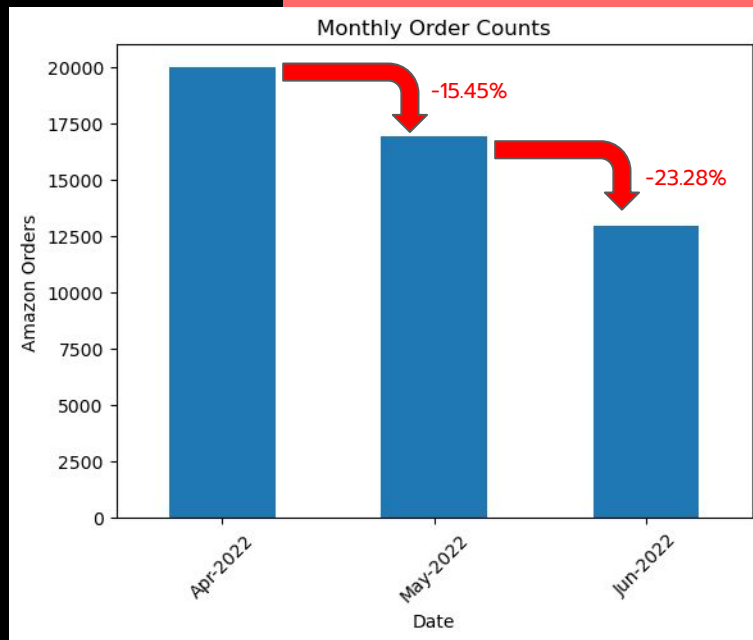
Difference in **average total orders** across months? 

- F-Statistic: 28.17
- P-value: 3.7e-10

Low p-value (> 0.05) and high F-stat indicate that there is a statistically significant relationship across average monthly orders.

See an increasing **decrease** in total monthly orders.

Month	
2022-04	19971
2022-05	16885
2022-06	12955

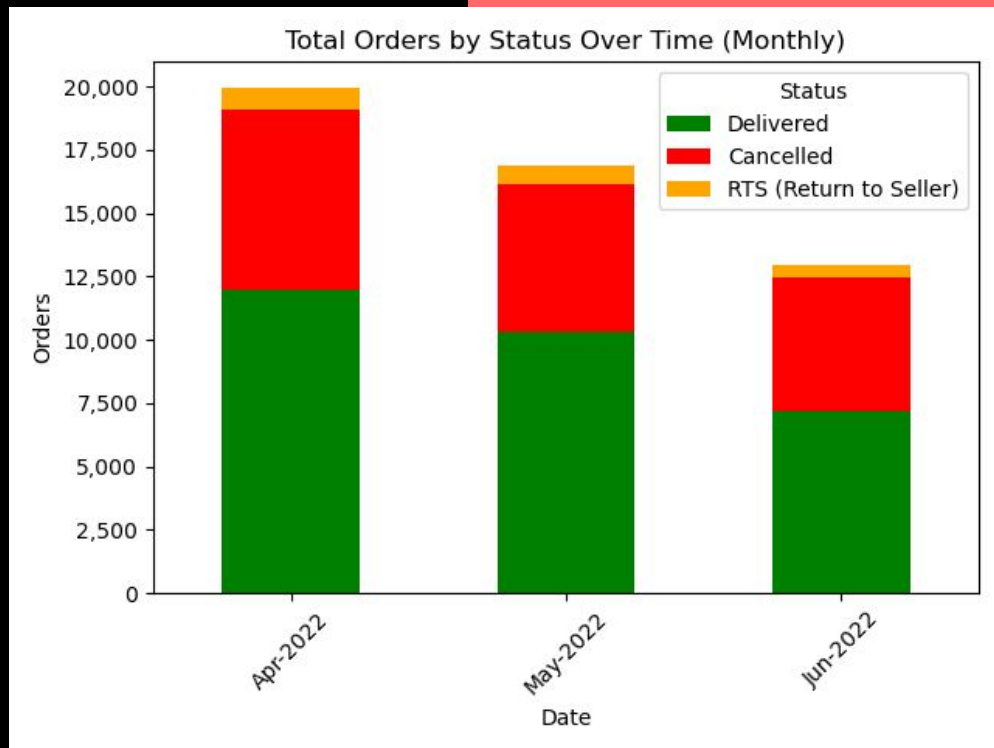


Monthly: Order Status

Relationship between
order status and
months? 

- Chi-square statistic: 139.63
- P-value: 3.39e-29

Very low p-value (< 0.05) and high chi-square statistic indicates statistically significant relationship between order status distribution (delivered, cancelled, or returned) and month.



Monthly: Order Status

April: Overall strong sales month but performed about as expected, however, did have 6.7% higher returns than expected.

Status	Total Orders	Success Rate (%)	Cancelled Rate (%)	Returned Rate (%)
Month				
2022-04	19971	59.891843	35.611637	4.496520
2022-05	16885	61.160794	34.604679	4.234528
2022-06	12955	55.453493	40.795060	3.751447

May: Strongest delivery performance with 5.5% less cancellations and 3.37% more successful deliveries than expected.

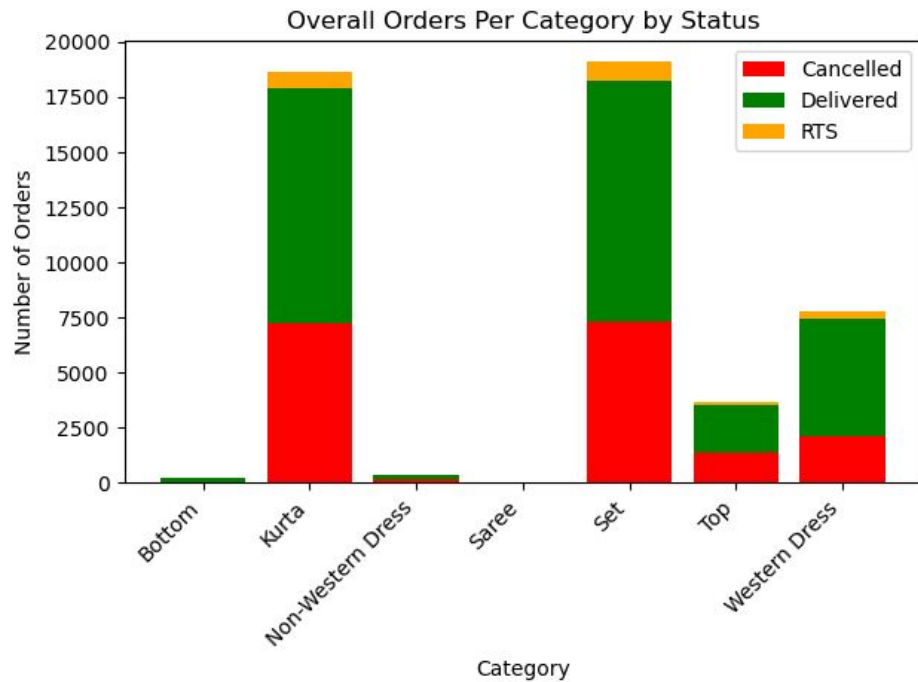
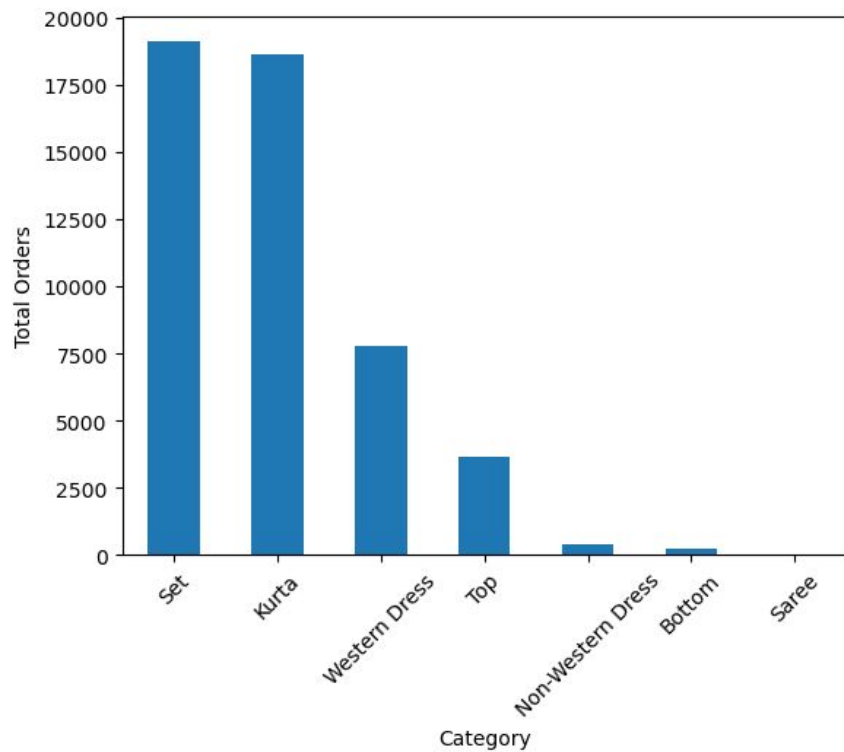
ORDER COUNT: OBSERVED - EXPECTED				
Status	Cancelled	Delivered	RTS (Return to Seller)	
Date				
2022-04-30	-201.064183	144.627873		56.436309
2022-05-31	-340.019815	336.541668		3.478147
2022-06-30	541.083998	-481.169541		-59.914457

June: Performed significantly worse than expected with 11.4% higher cancellations and 6.28% lower successful deliveries. But had 10.98% lower returns.

What are the most popular types of clothing items purchased in our data set?



Product Breakdown



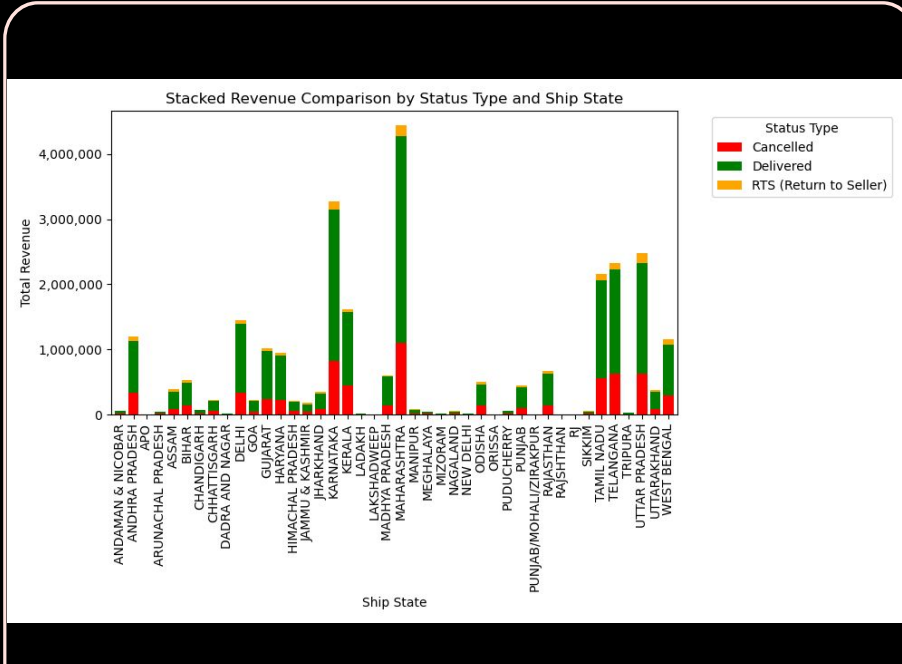
**How is revenue
reflected by what
state the items are
shipped to?**



Variables compared: status type, shipping state, and revenue.

Observations:

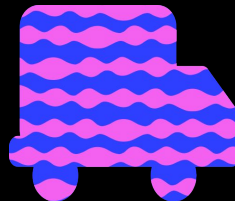
- **Top 3 states with the highest sales:**
 - **MAHARASHTRA = 3167712.0 INR**
 - **KARNATAKA = 2324568.0 INR**
 - **UTTAR PRADESH = 1693334.0 INR**
- **Bottom 3 states with the lowest sales:**
 - **RJ = 355.0 INR**
 - **PUNJAB/MOHALI/ZIRAKPUR = 568.0 INR**
 - **ORISSA = 1361.0 INR**



**What is the most
profitable shipping
service type?**



Variables compared: status type, shipping service type, and revenue.



Chi-square Statistic: 13.553823615822388

P-value: 0.00023181868181822973

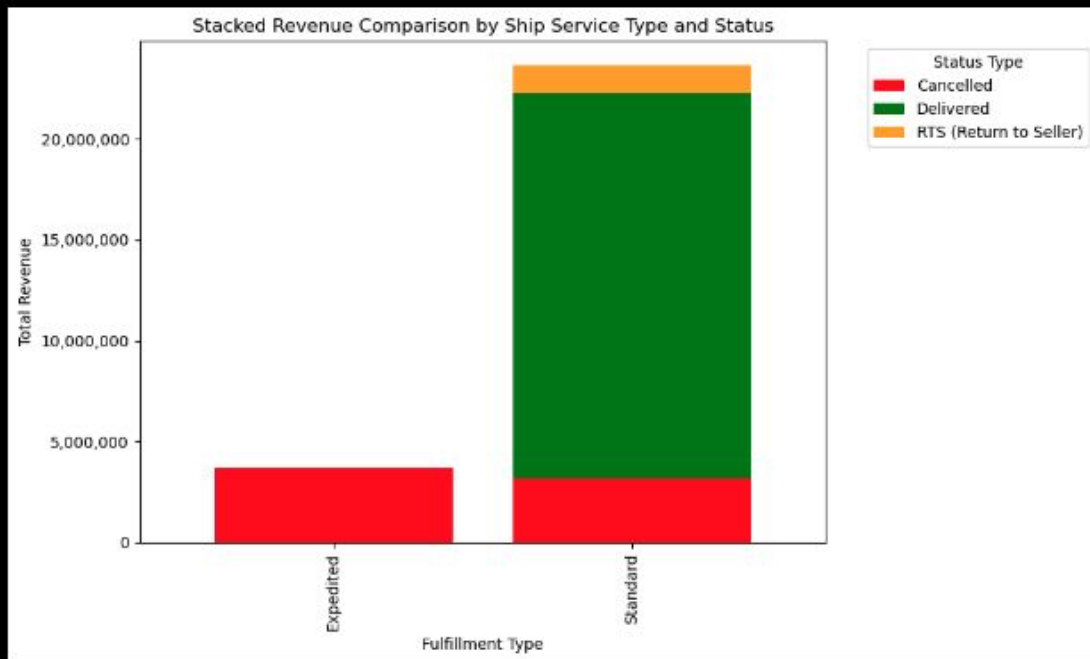
Degrees of Freedom: 1

Expected Frequencies:

[[5014.54290743 32162.45709257]

[579.45709257 3716.54290743]]

There is a significant relationship between revenue and ship-service-type.



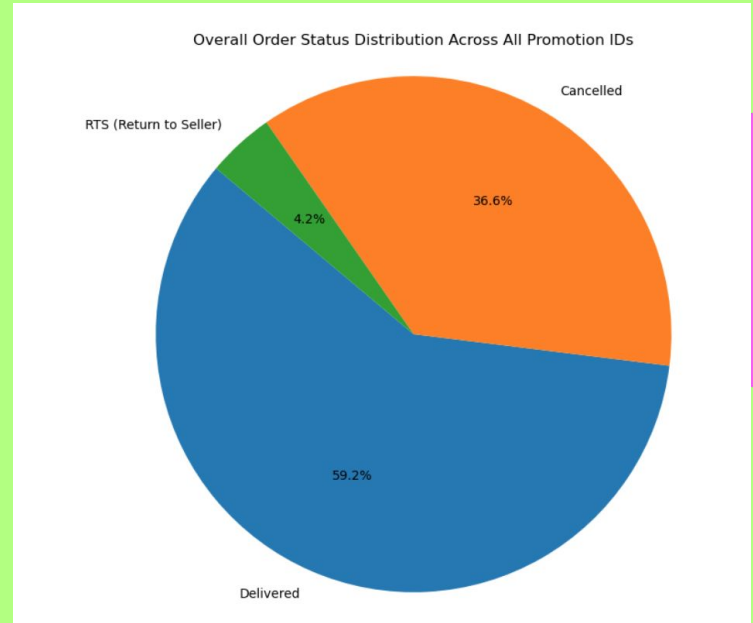
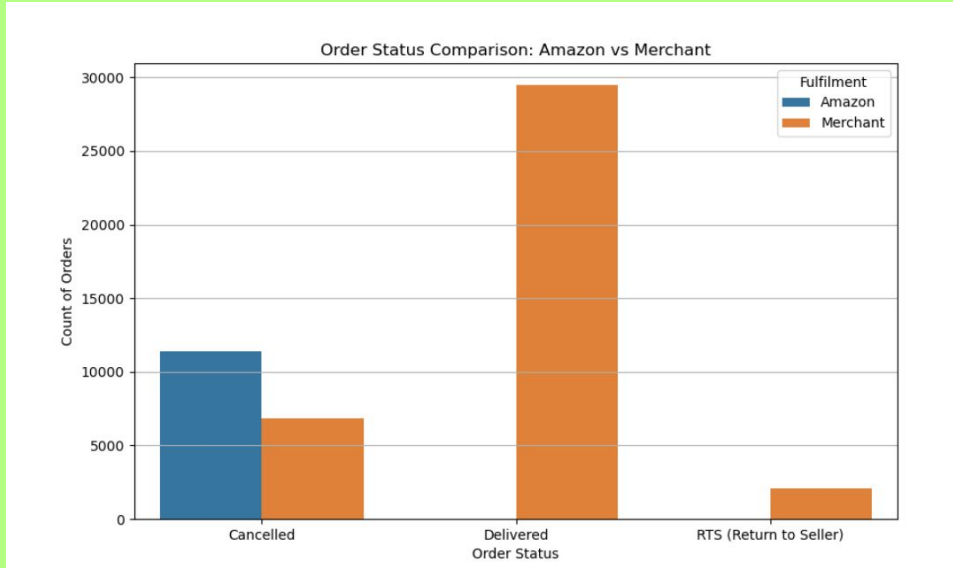
Observations:

- All of the orders that were expedited were cancelled.
- Standard delivery is the most common and the most profitable Shipping service type.
- So it may not be worth it to offer the expedited shipping option, for it is not as profitable for the seller
- Based on the chi test I did, it does appear that there is a significant relationship between revenue and ship service type.

**How do retailers
and promotions
performed in the
sales data
compare?**



- Data was grouped by "Fulfillment" and "Status," counting how many orders fell into each category.
- The majority of orders that were fulfilled by the merchant were successfully delivered.
- From the orders that were cancelled, Amazon had about 12,000 and the merchants had about 6,000.



- The pie chart provides a visual representation of how many orders fell into each status category across all promotion IDs. The orders that used the promotion id's came out to 59.2% delivered and 36.6% cancelled. Successful deliveries were more than half with orders using the promotion id's.

★ ADD TO CART ★

Retailers & Promotions

THANK YOU

