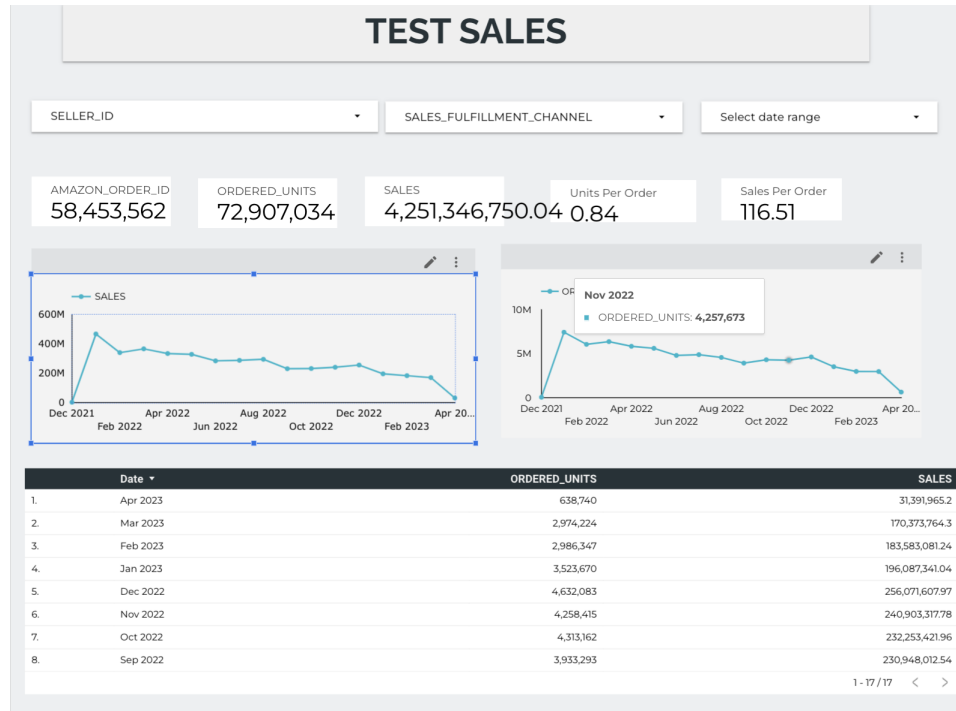


SPRINT 7 - SALES VALIDATION

1. Adding the DATE range filter in the dashboard



2. Validating for the users

List:

a40eb72f-b0ab-4db5-a068-46a6b01ae100
9de09329-e365-4dc9-8531-e74a23e8845a
0ea018e0-3f3e-4132-ae5b-cf119d480528
e6378394-c55c-434d-8ec9-74c2de63123f
dfaf3a9f-e94c-4315-b175-5eaf970c580d

- 0ea018e0-3f3e-4132-ae5b-cf119d480528

For scorecards adding the values for two years from the excel

Total Order Items	Units Ordered	Ordered Product Sales
614,676	678,121	\$12,915,040.12

<div> <div></div> <div></div> </div> <div> <div>AMAZON_ORDER_ID</div> <div>579,921</div> </div>	<div>ORDERED_UNITS</div> <div>681,882</div>	<div>SALES</div> <div>27,794,279.13</div>
---	---	---

Dashboard values , first two similar but the sales are nearly doubled. Diving deeper in monthly sales and values to recognize where the issues lie.

Observation:

For the first three months - the units are more in 2-3k. Which is fine and similar to others. But the sales are around double the value.

Looking for march sales to identify the issue.

For the query data which is on dashboard, column stats sum is

Sum	271,608.17 +
-----	--------------

Similar so okay.

Look at order 111-6087241-7493822 in whole data to see if any duplicate entries.

Not getting that in the actual data when adding filter in google sheets.

But there in data.

For other orders also somehow google sheets is not working.

Date - 1st march - in output table

111-6087241-7493822
111-8623348-6830645
112-4336489-2012239
112-6356739-9649856
113-4711001-3509831
113-7499737-0994603
114-0903677-6567438

But for the main table, we have many different amazon id's that are not present in the output query data at all.

Checking without the row number = 1 condition to see what is going wrong.

```
SELECT MP_SUP_KEY AS SELLER_ID ,purchase_day AS DATE_OF_PURCHASE, amazon_order_id AS
AMAZON_ORDER_ID, fulfillment_channel as SALES_FULFILLMENT_CHANNEL, ORDER_STATUS,
SUM(ITEM_PRICE * QUANTITY) OVER (PARTITION BY PURCHASE_DAY, MP_SUP_KEY,
amazon_order_id, partition_date ) AS SALES,
SUM(QUANTITY) OVER (PARTITION BY PURCHASE_DAY, MP_SUP_KEY , amazon_order_id,
partition_date ) AS ORDERED_UNITS,
ROW_NUMBER() OVER (PARTITION BY PURCHASE_DAY, MP_SUP_KEY, amazon_order_id) as rn
FROM `bigqueryexport-183608.amazon.customer_order_metrics`
WHERE EXTRACT(YEAR FROM purchase_date) IN (2022) AND EXTRACT(MONTH FROM
purchase_date) IN (3) AND ORDER_STATUS NOT IN ('Pending', 'Cancelled') AND MP_SUP_KEY
= '0ea018e0-3f3e-4132-ae5b-cf119d480528' ;
```

4	0ea018e0-3f3e-4132-ae5b-cf1...	2022-03-01	111-0009114-7170601	Merchant	Shipped	39.8	2	1
5	0ea018e0-3f3e-4132-ae5b-cf1...	2022-03-01	112-8591902-1962662	Amazon	Shipped	15.25	1	1
6	0ea018e0-3f3e-4132-ae5b-cf1...	2022-03-01	113-0488219-4738647	Amazon	Shipped	9.95	1	1
7	0ea018e0-3f3e-4132-ae5b-cf1...	2022-03-01	113-6067661-0974648	Amazon	Shipped	15.25	1	1
8	0ea018e0-3f3e-4132-ae5b-cf1...	2022-03-01	114-2991272-6709828	Amazon	Shipped	5.51	1	1
9	0ea018e0-3f3e-4132-ae5b-cf1...	2022-03-01	114-3556818-5800252	Amazon	Shipped	9.95	1	1
10	0ea018e0-3f3e-4132-ae5b-cf1...	2022-03-01	702-6252139-0137821	Amazon	Shipped	24.95	1	1
11	0ea018e0-3f3e-4132-ae5b-cf1...	2022-03-02	111-0702694-1019414	Amazon	Shipped	14.95	1	1
12	0ea018e0-3f3e-4132-ae5b-cf1...	2022-03-02	111-4446518-6073005	Merchant	Shipped	7.95	1	1

Again different order ids.



Some not showing. Let us order by date and explore with google sheets.

For same date getting 599 orders. Even the sum looks correct. 14,930 for both.

Observation + Realization:

For Jan itself there is a large error. Which is leading majorly in the scorecard error. SO now checking for Jan itself.

1. Duplicate data for one order:
Getting 20ish only on checking column stats. Trying to check in Google Sheets if unreasonably large value for any amazon_order_id (duplicates)?
No.
2. Which dates have maximum data. Let us check for those dates.

<div> <div>  Most </div> <div>  Least </div> </div>	
VALUE	FREQUENCY
1/13/2022	27,903
1/14/2022	23,128
1/15/2022	19,864
1/12/2022	19,754
1/17/2022	11,406

DATE_OF_PURCHASE	COUNTA of DATE_OF_PURCHASE
1/13/2022	27903
1/14/2022	23128
1/15/2022	19864
1/12/2022	19754
1/17/2022	11406
1/10/2022	11281
1/16/2022	9403
1/18/2022	8932
1/11/2022	8067
1/9/2022	7038

13th Jan check

Checking Sales column statistics.

Maximum value -

Total rows	27,903	+
Empty values	5	
Unique values	175	+
Sum	1,983,927.51	+
Average	71.11	+
Median	31.95	
Minimum value	0	+
Maximum value	79,875	+

Was able to fetch the mazon_orderR_id for 79875 . - 112-8533441-9757061

Checking records in data for this. Correct

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS		EXECUTION GRAPH		PREVIEW			
Row	amazon_order_id	order_status	item_price	quantity	last_updated_date	fulfillment_channel	ship_service_level	asin	sku	sal	
1	112-8533441-9757061	Shipped	1597.5	50	2022-01-16 17:12:50 UTC	Merchant	SecondDay	B08CSCKWM	FBM-ANFM-KN95-10pc	Am	

For value = 53955.

SELLER_ID	DATE_OF_P URCHASE	AMAZON_O RDER_ID	SALES_FUL FILLMENT_ CHANNEL	ORDER_STA TUS	SALES	ORDERED_U NITS
-----------	----------------------	---------------------	-----------------------------------	------------------	-------	-------------------

0ea018e0-3f3e-4132-ae5b-cf119d480528	1/13/2022	113-5940921-2628209	Merchant	Shipped	53955	30
--------------------------------------	-----------	---------------------	----------	---------	-------	----

Checking that order id.

```
SELECT * FROM `bigqueryexport-183608.amazon.customer_order_metrics`
WHERE EXTRACT(YEAR FROM purchase_date) IN (2022) AND EXTRACT(MONTH FROM purchase_date)
IN (1) AND ORDER_STATUS NOT IN ('Pending', 'Cancelled')
AND MP_SUP_KEY= '0ea018e0-3f3e-4132-ae5b-cf119d480528' and amazon_order_id=
'113-5940921-2628209';
```

CORRECT. Since the price value and quantity only was more.

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS		EXECUTION GRAPH		PREVIEW
Row	amazon_order_id	order_status	item_price	quantity	last_updated_date	fulfillment_channel	ship_service_level	
1	13-5940921-2628209	Shipped	1798.5	30	2022-01-15 09:08:44 UTC	Merchant	SecondDay	

FOR THE ENTIRE DATA:

Tr	Tr	Tr	Tr	Tr	123	123
SELLER_ID	DATE_OF	AMAZON_ORDER_ID	SALES_FL	ORDER_S	SALES	ORDERED
0ea018e0-3f3e-4	1/20/2022	114-6590725-7831464	Merchant	Shipped	7475000	500
0ea018e0-3f3e-4	1/11/2022	114-8284450-6813819	Merchant	Shipped	841200	200
0ea018e0-3f3e-4	1/21/2022	112-2071256-8837805	Merchant	Shipped	291375	150
0ea018e0-3f3e-4	1/20/2022	114-3791418-8208232	Merchant	Shipped	266895	90
0ea018e0-3f3e-4	1/21/2022	111-1605272-1284251	Amazon	Shipped	120500	100

Checking for 114-6590725-7831464

1 SELECT * FROM `bigqueryexport-183608.amazon.customer_order_metrics`
2 WHERE EXTRACT(YEAR FROM purchase_date) IN (2022) AND EXTRACT(MONTH FROM purchase_date) IN (
3 AND MP_SUP_KEY= '0ea018e0-3f3e-4132-ae5b-cf119d480528' and amazon_order_id= '114-6590725-78
4

Query results

JOB INFORMATIONRESULTSJSONEXECUTION DETAILSEXECUTION GRAPHPREVIEW

Row	amazon_order_id	order_status	item_price	quantity	last_up
1	114-6590725-7831464	Shipped	14950	500	2022-0

Also for entire data, for the same date getting same sales value with order prices but for different orders.
Why ?

Checking once.

0ea018e0-3f3e-4132-ae5b-cf119d480528	1/14/2022	114-7981876-4601015	Merchant	Shipped	35955	30
0ea018e0-3f3e-4132-ae5b-cf119d480528	1/14/2022	114-8071555-1236235	Merchant	Shipped	35955	30
0ea018e0-3f3e-4132-ae5b-cf119d480528	1/14/2022	112-1239833-1815447	Merchant	Shipped	35955	30
0ea018e0-3f3e-4132-ae5b-cf119d480528	1/14/2022	112-8716747-1433811	Merchant	Shipped	35955	30

Tt	📅	Tt	Tt	Tt	123	123
SELLER_ID	DATE_OF_	AMAZON_ORDER_ID	SALES_FL	ORDER_S	SALES	ORDERED
0ea018e0-3f3e-4132-ae5b-cf1	1/18/2022	111-7822872-3780217	Merchant	Shipped	55920	40
0ea018e0-3f3e-4132-ae5b-cf1	1/12/2022	113-5245956-6030650	Merchant	Shipped	53955	30
0ea018e0-3f3e-4132-ae5b-cf1	1/13/2022	113-5940921-2628209	Merchant	Shipped	53955	30
0ea018e0-3f3e-4132-ae5b-cf1	1/18/2022	111-5386470-2781862	Merchant	Shipped	52794.2	46
0ea018e0-3f3e-4132-ae5b-cf1	1/20/2022	114-6860135-3070633	Merchant	Shipped	52720	40
0ea018e0-3f3e-4132-ae5b-cf1	1/16/2022	114-7925711-7157011	Merchant	Shipped	44955	30
0ea018e0-3f3e-4132-ae5b-cf1	1/21/2022	111-2603194-6565869	Amazon	Shipped	37375	50
0ea018e0-3f3e-4132-ae5b-cf1	1/14/2022	112-5312648-2134632	Merchant	Shipped	35955	30
0ea018e0-3f3e-4132-ae5b-cf1	1/14/2022	114-7981876-4601015	Merchant	Shipped	35955	30
0ea018e0-3f3e-4132-ae5b-cf1	1/14/2022	114-8071555-1236235	Merchant	Shipped	35955	30
0ea018e0-3f3e-4132-ae5b-cf1	1/14/2022	112-1239833-1815447	Merchant	Shipped	35955	30
0ea018e0-3f3e-4132-ae5b-cf1	1/14/2022	112-8716747-1433811	Merchant	Shipped	35955	30
0ea018e0-3f3e-4132-ae5b-cf1	1/19/2022	114-6765594-6077860	Merchant	Shipped	31455	30
0ea018e0-3f3e-4132-ae5b-cf1	1/27/2022	111-4644555-1822627	Amazon	Shipped	23283	66
0ea018e0-3f3e-4132-ae5b-cf1	1/17/2022	114-1636554-7666608	Merchant	Shipped	19980	20
0ea018e0-3f3e-4132-ae5b-cf1	1/13/2022	113-2639678-4530661	Merchant	Shipped	19968.75	25
0ea018e0-3f3e-4132-ae5b-cf1	1/16/2022	111-8352092-8958603	Merchant	Shipped	19775	30
0ea018e0-3f3e-4132-ae5b-cf1	1/14/2022	113-8549665-5494618	Merchant	Shipped	15980	20
0ea018e0-3f3e-4132-ae5b-cf1	1/13/2022	112-7001776-3673816	Merchant	Shipped	15980	20
0ea018e0-3f3e-4132-ae5b-cf1	1/14/2022	111-0656666-7545021	Merchant	Shipped	15980	20
0ea018e0-3f3e-4132-ae5b-cf1	1/13/2022	114-6000086-4229040	Merchant	Shipped	15980	20
0ea018e0-3f3e-4132-ae5b-cf1	1/14/2022	112-5069070-9937838	Merchant	Shipped	15980	20
0ea018e0-3f3e-4132-ae5b-cf1	1/18/2022	113-4071591-1505811	Merchant	Shipped	15593.75	25
0ea018e0-3f3e-4132-ae5b-cf1	1/18/2022	111-8371136-4441851	Merchant	Shipped	15593.75	25
0ea018e0-3f3e-4132-ae5b-cf1	1/19/2022	114-1497177-1185009	Merchant	Shipped	13980	20

SAME Seller, SAME Date, SAME Sales, SAME ordered units, DIFFERENT **AMAZON_ORDER_ID's**.
Strangely multiple instances in the entire month of **JANUARY**.

For those orders, ship service label is usually different/varying between two options. IS IT LEADING TO THIS ERROR? Business context of ORDER_ID ?

```
SELECT * FROM `bigqueryexport-183608.amazon.customer_order_metrics`
WHERE EXTRACT(YEAR FROM purchase_date) IN (2022) AND EXTRACT(MONTH FROM purchase_date)
IN (1) AND ORDER_STATUS NOT IN ('Pending', 'Cancelled')
AND MP_SUP_KEY= '0ea018e0-3f3e-4132-ae5b-cf119d480528'
AND amazon_order_id IN ('114-7981876-4601015', '114-8071555-1236235',
'112-1239833-1815447', '112-8716747-1433811') ;
```


JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS		EXECUTION GRAPH		PREVIEW			
Row	amazon_order_id	order_status	item_price	quantity	last_updated_date	fulfillment_channel	ship_service_level	asin	sku	sales_channel	
1	114-7981876-4601015	Shipped	1198.5	30	2022-01-16 03:14:38 UTC	Merchant	SecondDay	B08BF3W77K	FBM-ANFM-KN95-50pc	Amazon.com	
2	112-1239833-1815447	Shipped	1198.5	30	2022-01-16 03:33:37 UTC	Merchant	SecondDay	B08BF3W77K	FBM-ANFM-SURGL-8	Amazon.com	
3	112-8716747-1433811	Shipped	1198.5	30	2022-01-16 03:13:35 UTC	Merchant	SecondDay	B08BF3W77K	FBM-ANFM-SURGL-8	Amazon.com	
4	114-8071555-1286235	Shipped	1198.5	30	2022-01-15 08:59:06 UTC	Merchant	NextDay	B08BF3W77K	FBM-ANFM-KN95-50pc	Amazon.com	

Seller Data VALIDATION FOR : dfaf3a9f-e94c-4315-b175-5eaf970c580d

Total Order Items

Units Ordered

Ordered Product Sales

478,793

521,903

\$18,552,686.75

AMAZON_ORDER_ID	ORDERED_UNITS	SALES
399,023	466,984	19,708,843.28

1 M over a period of 1.5 years is a decent difference of error.

Even checked on a monthly basis, looks good. Difference of around 100k each month for 1.2M sales.

JOINING COMPETITIVE PRICING WITH THE LISTINGS TABLE:

I want to see that if I join both the tables on basis of the ASIN value, is there a point where the listing price from the listings table matches with the new_buy_box_listing price?

If it ever does match, it may mean that that seller was WON THE BUY BOX.

```
SELECT open_date, YEAR, LISTING_TABLE_PRICE, SELLER_ID, CP.asin AS ASIN,
CP.new_buy_box_listing_price_amount ,CP.listing_count_offers, CP.path_golden
FROM
(SELECT SAFE_CAST(LEFT(open_date,10) AS DATE) AS open_date, REGEXP_EXTRACT(open_date,
r'\d{4}') AS YEAR, price as LISTING_TABLE_PRICE, mp_sup_key AS SELLER_ID, ASIN AS ASIN
FROM `bigqueryexport-183608.amazon.listings`
WHERE REGEXP_EXTRACT(open_date, r'\d{4}') IN ('2022', '2023')) L
INNER JOIN `bigqueryexport-183608.amazon.competitive_pricing` CP
ON L.ASIN = CP.ASIN
WHERE CP.ASIN= 'B08TTS7VP3';
```

LISTING_TABLE_PRICE	SELLER_ID	ASIN	new_buy_box_listing_price_amount	listing_count_offers
23.99	63959617-ac0c-42ce-9514-2ef...	B08TTS7VP3	25.44	71
23.99	63959617-ac0c-42ce-9514-2ef...	B08TTS7VP3	14.19	2
23.99	63959617-ac0c-42ce-9514-2ef...	B08TTS7VP3	15.99	9
23.99	63959617-ac0c-42ce-9514-2ef...	B08TTS7VP3	25.44	71
23.99	63959617-ac0c-42ce-9514-2ef...	B08TTS7VP3	19.99	178

Does not make sense. Checked for a few asins.

Not the correct approach.

It is not possible to get the seller information too.