**Myntra ad hoc analysis project report**

**Abhishek Sharma**

Based on products listed till 4TH MAY 2023

Project report

**Q1. What is the average price of products listed on Myntra?**

* 1538.2

*SQL CODE*

* select avg(price) from myntra\_selected

**Q2. How many different sellers are there on Myntra?**

* 5482

*SQL CODE*

* select count (distinct

seller )

from

myntra\_selected

**Q3. What are the top five sellers on Myntra based on the number of products they offer? Provide both the seller name and the number of products they sell.**

|  |  |
| --- | --- |
| **seller** | **no\_of\_products** |
| Roadster | 10594 |
| H&M | 6649 |
| Puma | 6525 |
| max | 6457 |
| Anouk | 6078 |

|  |  |
| --- | --- |
| **seller** | **no\_of\_products** |
| Wild & Free | 1 |
| Reistor | 1 |
| Silverholic | 1 |
| Amefa | 1 |
| TRU HAIR | 1 |

* *SQL CODE*

select seller,

count(\*) as no\_of\_products

from myntra\_selected

group by seller

order by count(\*) desc

limit 5

**Q4. What is the average rating of products listed on Myntra?**

* 1.60
* *SQL CODE*

select avg(rating) from myntra\_selected

**Q5. List the names of T-shirts priced higher than the average T-shirt price.**

|  |
| --- |
| Men Striped Polo Collar T-shirt |
| Men Polo Collar Slim Fit T-shirt |
| Men Printed Football T-shirt |
| Men Solid Pure Cotton T-shirt |
| Men Solid Pure Cotton T-shirt |

* *SQL CODE*

with cte as (SELECT \*,

case when name like '%Men%' then name

end as men ,

CASE

WHEN name LIKE '%T-shirt%' THEN name

END AS "T-shirts"

FROM myntra\_selected

WHERE CASE

WHEN name LIKE '%T-shirt%' THEN name

END IS NOT NULL

and

case when name like '%Men%' then name

end is not null

)

select \* from cte

where price > (select avg(price) from cte)

order by price desc

limit 5

**Q6. How many products are available for men, women, boys, and girls respectively?**

|  |  |  |  |
| --- | --- | --- | --- |
| **total\_men\_products** | **total\_women\_products** | **total\_boys\_products** | **total\_girls\_products** |
| 200007 | 145146 | 41649 | 16096 |

* SQL CODE

with cte as (select \*,

CASE

WHEN name LIKE '%Men%' THEN name

END AS men,

CASE

WHEN name LIKE '%Women%' THEN name

END AS Women,

CASE

WHEN name LIKE '%Boys%' THEN name

END AS Boys,

CASE

WHEN name LIKE '%Girls%' THEN name

END AS Girls

from myntra\_selected)

select count (\*)as total\_men\_Products,

(select count(\*) from cte where Women is not null ) as Total\_Women\_Products,

(select count(\*) from cte where Boys is not null ) as Total\_Boys\_Products,

(select count(\*) from cte where Girls is not null ) as Total\_Girls\_Products

from cte where men is not null

**Q7. What are the average prices of products and discounts offered for men's and women's items?**

|  |  |
| --- | --- |
| **avg\_price\_men** | **avg\_price\_women** |
| 1525.1 | 1644.76 |

* SQL CODE

with cte as (select \*,

CASE

WHEN name LIKE '%Men%' THEN name

END AS men ,

CASE

WHEN name LIKE '%Women%' THEN name

END AS Women from myntra\_selected)

select round (avg(price),2) as Avg\_price\_men,

(select round (avg(price),2) as Avg\_price\_Women

from cte

where Women is not null)

from cte

where men is not null

**Q8. Compare the average discount percentage offered for men's, women's, and kids' items.**

|  |  |
| --- | --- |
| **avg\_discount\_men** | **avg\_discount\_women** |
| 92.69 | 120.16 |

* SQL CODE

with cte as (select \*,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct ,

CASE

WHEN name LIKE '%Men%' THEN name

END AS men ,

CASE

WHEN name LIKE '%Women%' THEN name end as Women

from myntra\_selected)

select round (avg(discount),2) as Avg\_price\_men,

(select round (avg(discount),2) as Avg\_price\_Women

from cte

where Women is not null)

from cte

where men is not null

|  |  |
| --- | --- |
| **avg\_rating\_mens\_products** | **avg\_rating\_womens\_products** |
| 1.44 | 1.24 |

**Q9. What is the average rating for men's and women's products respectively?**

with cte as (select \*,

CASE

WHEN name LIKE '%Men%' THEN name

END AS men ,

CASE

WHEN name LIKE '%Women%' THEN name end as Women

from myntra\_selected)

select avg(rating) as Avg\_rating\_mens\_products,

(select avg(rating)as Avg\_rating\_Womens\_products

from cte

where Women is not null)

from cte

where men is not null

**Q10. Who are the top five sellers on Myntra based on the average rating and the total number of ratings received?**

|  |  |  |
| --- | --- | --- |
| **seller** | **avg\_rating** | **no\_of\_reviews** |
| Biotique | 4 | 13240.13 |
| Donald Duck | 4 | 5600 |
| Neutrogena | 4 | 5500.3 |
| Maybelline | 4 | 5328.39 |
| indulekha | 4 | 3955.18 |

* SQL CODE

Select seller,

round (avg(rating)::int,2) as avg\_rating,

round (avg(ratingTotal),2) as No\_of\_reviews

from myntra\_selected

group by seller

order by round (avg(ratingTotal),2) desc , round (avg(rating)::int,2) desc

limit 5

**Q11. Which brands offer the highest average discount? Provide separate values for men's and women's items.**

|  |  |
| --- | --- |
| **avg\_discount\_men** | **avg\_discount\_women** |
| 39.12 | 42.52 |
|  |  |

* SQL CODE

with cte as (select \*,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct ,

CASE

WHEN name LIKE '%Men%' THEN name

END AS men ,

CASE

WHEN name LIKE '%Women%' THEN name end as Women

from myntra\_selected)

select round (avg(discount\_prct),2) as Avg\_price\_men,

(select round (avg(discount\_prct),2)

from cte

where Women is not null) as Avg\_price\_Women

from cte

where men is not null

With cte as (with cte as (select \*,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct ,

CASE

WHEN name LIKE '%Men%' THEN name

END AS men ,

CASE

WHEN name LIKE '%Women%' THEN name end as Women

from myntra\_selected)

select round (avg(discount\_prct),2) as Avg\_discount\_men,

(select round (avg(discount\_prct),2)

from cte

where Women is not null) as Avg\_discount\_Women

from cte

where men is not null ),

men as (select \* ,

(select avg\_discount\_men from cte) as avg\_discount\_men,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct

from myntra\_selected

where

CASE

WHEN name LIKE '%Men%' THEN name

END is not null ),

women as (select \* ,

(select avg\_discount\_women from cte) as avg\_discount\_women,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct

from myntra\_selected

where

CASE

WHEN name LIKE '%Women%' THEN name

END is not null )

select

name,

avg\_discount\_men,

discount\_prct,

(discount\_prct- avg\_discount\_men) as discount\_difference

from men

where discount\_prct is not null

order by (discount\_prct- avg\_discount\_men) desc

limit 5

**Q 10. Best sellers based on avg rating, no of ratings**

Q10. Which brands offer most discount?

|  |  |  |  |
| --- | --- | --- | --- |
| **name** | **avg\_discount\_men** | **discount\_prct** | **discount\_difference** |
| Men Penis Enlargement Oil | 39.12 | 96 | 56.88 |
| Men Printed SPTCAS T-shirt | 39.12 | 96 | 56.88 |
| Men Mid-Rise Joggers | 39.12 | 96 | 56.88 |
| Speedy & Easy Men's Hair Color | 39.12 | 96 | 56.88 |
| Men Printed Trunk | 39.12 | 96 | 56.88 |

With cte as (with cte as (select \*,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct ,

CASE

WHEN name LIKE '%Men%' THEN name

END AS men ,

CASE

WHEN name LIKE '%Women%' THEN name end as Women

from myntra\_selected)

select round (avg(discount\_prct),2) as Avg\_discount\_men,

(select round (avg(discount\_prct),2)

from cte

where Women is not null) as Avg\_discount\_Women

from cte

where men is not null ),

men as (select \* ,

(select avg\_discount\_men from cte) as avg\_discount\_men,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct

from myntra\_selected

where

CASE

WHEN name LIKE '%Men%' THEN name

END is not null ),

women as (select \* ,

(select avg\_discount\_women from cte) as avg\_discount\_women,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct

from myntra\_selected

where

CASE

WHEN name LIKE '%Women%' THEN name

END is not null )

select

name,

avg\_discount\_women,

discount\_prct,

(discount\_prct- avg\_discount\_women) as discount\_difference

from women

where discount\_prct is not null

group by name,avg\_discount\_women,discount\_prct,(discount\_prct- avg\_discount\_women)

order by (discount\_prct- avg\_discount\_women) desc

limit 5

|  |  |  |  |
| --- | --- | --- | --- |
| **name** | **avg\_discount\_women** | **discount\_prct** | **discount\_difference** |
| Women Pack Of 3 Bikini Briefs | 42.52 | 96 | 53.48 |
| Women Pack Of 3 Hipster Briefs | 42.52 | 96 | 53.48 |
| Women Pack Of 3 Printed Bikini Briefs | 42.52 | 96 | 53.48 |
| Women Pack Of 3 Printed Cotton Bikini Briefs | 42.52 | 96 | 53.48 |
| Women Pack Of 3 Solid Bikini Briefs | 42.52 | 96 | 53.48 |

So we can conclude that in both men and women undergarments is higly discounted item.

**Q12. Identify the products with the highest discount compared to the category average. Provide the product name, category, and the difference in discount percentage.**

|  |  |  |  |
| --- | --- | --- | --- |
| **category** | **total\_rating** | **avg\_rating** | **no\_of\_sellers** |
| Pure Cotton T-shirt | 468688 | 2 | 217 |
| Slim Fit Jeans | 412097 | 2 | 153 |
| Skinny Fit Jeans | 325245 | 2 | 92 |
| Polo Collar T-shirt | 248376 | 2 | 314 |
| Tapered Fit Jeans | 232374 | 2 | 54 |

|  |  |  |  |
| --- | --- | --- | --- |
| **category** | **total\_rating** | **avg\_rating** | **no\_of\_sellers** |
| Women Analogue Watch | 268092 | 2 | 122 |
| NULL | 81290 | 1 | 635 |
| Pure Cotton T-shirt | 78242 | 3 | 49 |
| Women Walking Shoes | 48555 | 2 | 51 |
| Women Running Shoes | 48172 | 2 | 59 |

with cte as (with cte as (select \*,

CASE WHEN name LIKE '%Women%' THEN name END AS Women

from myntra\_selected

where

CASE WHEN name LIKE '%Women%' THEN name END is not null )

select \*,

STRING\_TO\_ARRAY(name, ' ') AS name\_array

from cte

order by ratingtotal desc, rating desc )

select

name\_array[array\_length(name\_array,1)-2]||' ' || name\_array[array\_length(name\_array,1)-1] || ' ' || name\_array[array\_length(name\_array,1)] as category ,

sum(ratingtotal) as Total\_rating,

avg(rating)::int as Avg\_rating,

count(distinct seller) as No\_of\_Sellers

from cte

group by name\_array[array\_length(name\_array,1)-2]||' ' || name\_array[array\_length(name\_array,1)-1] || ' ' || name\_array[array\_length(name\_array,1)]

order by sum(ratingtotal) desc , avg(rating)::int desc

limit 5

with cte as (with cte as (select \*,

CASE WHEN name LIKE '%Men%' THEN name END AS Men

from myntra\_selected

where

CASE WHEN name LIKE '%Men%' THEN name END is not null )

select \*,

STRING\_TO\_ARRAY(name, ' ') AS name\_array

from cte

order by ratingtotal desc, rating desc )

select

name\_array[array\_length(name\_array,1)-2]||' ' || name\_array[array\_length(name\_array,1)-1] || ' ' || name\_array[array\_length(name\_array,1)] as category ,

sum(ratingtotal) as Total\_rating,

avg(rating)::int as Avg\_rating,

count(distinct seller) as No\_of\_Sellers

from cte

group by name\_array[array\_length(name\_array,1)-2]||' ' || name\_array[array\_length(name\_array,1)-1] || ' ' || name\_array[array\_length(name\_array,1)]

order by sum(ratingtotal) desc , avg(rating)::int desc

limit 5

**Q13. Determine the categories and sellers with the most products rated below the category average. Provide the category name, total products, total low-rated products, and the seller name.**

|  |  |  |
| --- | --- | --- |
| **category** | **total\_products** | **total\_low\_rating\_products** |
| Pure Cotton T-shirt | 18210 | 7605 |
| Polo Collar T-shirt | 17414 | 9013 |
| Slim Fit T-shirt | 15146 | 9073 |
| 2 Pillow Covers | 10820 | 8615 |
| Printed Cotton T-shirt | 7997 | 6912 |

|  |  |  |  |
| --- | --- | --- | --- |
| **category** | **seller** | **total\_products** | **total\_low\_rating\_products** |
| 2 Pillow Covers | JC Collection | 747 | 741 |
| 2 Pillow Covers | SWAYAM | 566 | 436 |
| 2 Lounge Pants | FashionRack | 561 | 550 |
| Slim Fit T-shirt | U.S. Polo Assn. | 537 | 83 |
| Gold-Plated Jewellery Set | ROFARWORD | 474 | 472 |

SQL CODE

with cte as (select \*,

STRING\_TO\_ARRAY(name, ' ') AS name\_array

from myntra\_selected

order by ratingtotal desc, rating desc)

, cte2 as (

select

name\_array[array\_length(name\_array,1)-2]||' ' || name\_array[array\_length(name\_array,1)-1] || ' ' || name\_array[array\_length(name\_array,1)] as category ,

sum(ratingtotal) as Total\_rating,

avg(rating)::int as Avg\_rating,

count(distinct seller) as No\_of\_Sellers

from cte

group by name\_array[array\_length(name\_array,1)-2]||' ' || name\_array[array\_length(name\_array,1)-1] || ' ' || name\_array[array\_length(name\_array,1)]

order by sum(ratingtotal) desc , avg(rating)::int desc ),

cte3 as (select \*,

name\_array[array\_length(name\_array,1)-2]||' ' || name\_array[array\_length(name\_array,1)-1] || ' ' || name\_array[array\_length(name\_array,1)] as category

from

cte),

cte4 as (

select cte3.category,

count(\*) as no\_of\_products

from cte3

join cte2

on cte3.category= cte2.category

where

rating <avg\_rating

group by cte3.category

order by count(\*) desc ),

cte5 as (select cte3.category,

count(cte3.category) as Total\_products

from cte3

group by cte3.category

order by count(\*) desc )

select cte4.category,

cte5.total\_products,

cte4.no\_of\_products as Total\_low\_rating\_Products

from cte5 join cte4 on

cte4.category= cte5.category

with cte as (select \*,

STRING\_TO\_ARRAY(name, ' ') AS name\_array

from myntra\_selected

order by ratingtotal desc, rating desc)

, cte2 as (

select

name\_array[array\_length(name\_array,1)-2]||' ' || name\_array[array\_length(name\_array,1)-1] || ' ' || name\_array[array\_length(name\_array,1)] as category ,

sum(ratingtotal) as Total\_rating,

avg(rating)::int as Avg\_rating,

count(distinct seller) as No\_of\_Sellers

from cte

group by name\_array[array\_length(name\_array,1)-2]||' ' || name\_array[array\_length(name\_array,1)-1] || ' ' || name\_array[array\_length(name\_array,1)]

order by sum(ratingtotal) desc , avg(rating)::int desc ),

cte3 as (select \*,

name\_array[array\_length(name\_array,1)-2]||' ' || name\_array[array\_length(name\_array,1)-1] || ' ' || name\_array[array\_length(name\_array,1)] as category

from

cte),

cte4 as (

select cte3.category,

seller,

count(\*) as no\_of\_products

from cte3

join cte2

on cte3.category= cte2.category

where

rating <avg\_rating

group by cte3.category, seller

order by count(\*) desc ),

cte5 as (select cte3.category,

seller,

count(cte3.category) as Total\_products

from cte3

group by cte3.category , seller

order by count(\*) desc )

select cte4.category,

cte5.seller,

cte5.total\_products,

cte4.no\_of\_products as Total\_low\_rating\_Products

from cte5 join cte4 on

cte4.category= cte5.category and cte5.seller=cte4.seller

limit 5

Q14. What is the average discount percentage for men's shoes on Myntra?

36.88

SQL CODE

select round (avg(discount\_prct),2) as discount\_prct

from

(select \*,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct,

STRING\_TO\_ARRAY(name, ' ') as name\_array

from myntra\_selected

where name like '%Shoes%'

or name like '%shoes%') f

where

name\_array[1]='Men'

Q15. Compare the average discount percentages for men's, women's, and kids' shoes respectively.

|  |  |  |
| --- | --- | --- |
| **avg\_discount\_kids\_shoes** | **avg\_discount\_men\_shoes** | **avg\_discount\_women\_shoes** |
| 34.1 | 36.88 | 30.81 |

* SQL CODE

with men as (select round (avg(discount\_prct),2) as discount\_prct

from

(select \*,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct,

STRING\_TO\_ARRAY(name, ' ') as name\_array

from myntra\_selected

where name like '%Shoes%'

or name like '%shoes%') f

where

name\_array[1]='Men')

, women as (select round (avg(discount\_prct),2) as discount\_prct

from

(select \*,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct,

STRING\_TO\_ARRAY(name, ' ') as name\_array

from myntra\_selected

where name like '%Shoes%'

or name like '%shoes%') f

where

name\_array[1]='Women'),

kids as (select round (avg(discount\_prct),2) as discount\_prct

from

(select \*,

case

when discount >=100 then round (((mrp- price)/mrp)\*100 ,2)

when discount <99 then discount

end as discount\_prct,

STRING\_TO\_ARRAY(name, ' ') as name\_array

from myntra\_selected

where name like '%Shoes%'

or name like '%shoes%') f

where

name\_array[1]='Kids')

select discount\_prct as Avg\_discount\_kids\_shoes ,

(select \* from men) as Avg\_discount\_Men\_shoes,

(select \* from women) as Avg\_discount\_Women\_shoes

from kids