SQL - Shark Tank India Project

Dipti Soni

Introduction

Shark Tank India is a reality TV show where entrepreneurs pitch their business ideas to a panel of investors, called "Sharks".

The Sharks are seasoned business experts who decide whether to invest in the entrepreneurs' business ideas. If they do invest, they may also help to market the product and mentor the entrepreneurs.

Aspiring entrepreneurs from India pitch their business models to a panel of investors and persuade them to invest money in their idea.

Objective

3 Seasons of Shark Tank India are over.

Based on this data, we will derive some insights for future Shark Tank India Seasons.

sharkname	average
Namita	33.71
Vineeta	31.25
Anupam	29.99
Aman	34.89

season_number	total	r_offer	a_offer	r_offer/total*100	a_offer/total*100
1	304	192	140	63.1579	46.0526
2	338	242	212	71.5976	62.7219
3	314	208	184	66.2420	58.5987

Industry	Total_Deal_Amount_in_lakhs
Agriculture	75
Animal/Pets	60
Beauty/Fashion	300

Data Cleaning Process

The dataset has only one table with huge number of columns - 52

First the data is cleaned using pandas. The numbers of columns are reduced from 52 to 42.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 478 entries, 0 to 477
Data columns (total 52 columns):
 #   Column                              Non-Null Count  Dtype
---  ------                              --------------  -----
 0   Season_Number                       478 non-null    int64
 1   Startup_Name                        478 non-null    object
 2   Episode_Number                      478 non-null    int64
```

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 478 entries, 0 to 477
Data columns (total 42 columns):
 #   Column                              Non-Null Count  Dtype
---  ------                              --------------  -----
 0   Season_Number                       478 non-null    int64
 1   Startup_Name                        478 non-null    object
 2   Episode Number                      478 non-null    int64
```

Highest Funding by Domain

Your Team must promote shark Tank India season 4, a senior comes up with an idea to show highest funding given till date for each domain so that new startups can be attracted, and you were assigned the task to show the same

```
select * from
(select Industry, Total_Deal_Amount_in_lakhs,
ROW_NUMBER() over(PARTITION BY industry order by Total_Deal_Amount_in_lakhs desc) as 'ranking' from sharktank)t
where ranking = 1;
```

	Industry	Total_Deal_Amount_in_lakhs	ranking
▶	Agriculture	75	1
	Animal/Pets	60	1
	Beauty/Fashion	300	1
	Education	150	1

Female to Male Pitcher Ratio by Domain

You have been assigned the role of finding the domain where female to male pitcher ratio > 70%

```
select industry, round(sum(female_presenters)/sum(male_presenters)*100,2) as ratio from sharktank
group by industry
having ratio > 70;
```

industry	ratio
Beauty/Fashion	85.56
Education	73.33
Animal/Pets	71.43

Pitch Volume and Conversion Rate

You are working at marketing firm of Shark Tank India, you have got the task to determine volume of per season sale pitch made, pitches that received offer and pitches that were converted. Also show the percentage of pitches converted and percentage of pitches entertained.

```
select a.season_number,total, r_offer, a_offer, r_offer/total*100, a_offer/total*100 from
(select season_number, count(pitch_number) as 'total' from sharktank group by season_number) a
inner join
(select season_number, received_offer, count(pitch_number) as 'r_offer' from sharktank where received_offer = 'Yes' group by season_number) b
on a.season_number = b.season_number
inner join
(select season_number, accepted_offer, count(pitch_number) as 'a_offer' from sharktank where accepted_offer = 'Yes' group by season_number) c
on b.season_number = c.season_number ;
```

season_number	total	r_offer	a_offer	r_offer/total*100	a_offer/total*100
1	304	192	140	63.1579	46.0526
2	338	242	212	71.5976	62.7219
3	314	208	184	66.2420	58.5987

Season with the highest monthly sales for industries

As a venture capital firm specializing in investing in startups featured on a renowned entrepreneurship TV show, you are determining the season with the highest average monthly sales and identify the top 5 industries with the highest average monthly sales during that season to optimize investment decisions?

```
set @seas= (select season_number  from
(
select  season_number , round(avg(monthly_sales_in_lakhs),2)as 'average'
from sharktank
where monthly_sales_in_lakhs!= 'Not_mentioned'
group by season_number
)k order by average desc
limit 1);

select industry , round(avg(monthly_sales_in_lakhs),2) as average
from sharktank
where season_number = @seas and monthly_sales_in_lakhs!= 'Not_mentioned'
group by industry
order by average desc
limit 5;
```

industry	average
Electronics	3500
Furnishing/Household	96.27
Beauty/Fashion	61.65
Food	41.28
Services	29

Industries with consistent increase in funding

As a data scientist at our firm, your role involves solving real-world challenges like identifying industries with consistent increase in funds raised over multiple seasons. This requires focusing on industries where data is available across all three years. Once these industries are pinpointed, your task is to delve into the specifics, analyzing the number of pitches made, offers received, and offers converted per season within each industry.

```
with validindustries as (
select industry,
sum(case when season_number = 1 then total_deal_amount_in_lakhs end) as season1,
sum(case when season_number = 2 then total_deal_amount_in_lakhs end) as season2,
sum(case when season_number = 3 then total_deal_amount_in_lakhs end) as season3
from sharktank
group by industry having season3 > season2 and season2 > season1 and season1 != 0)

select v.industry, s.season_number,
count(s.startup_name) as 'no_of_startups', count(case when s.Received_Offer='Yes' then Received_Offer end) as 'received',
count(case when s.Accepted_Offer='Yes' then Accepted_Offer end) as 'accepted'
from sharktank s join validindustries v where s.industry = v.industry
group by industry, season_number;
```

industry	season_number	no_of_startups	received	accepted
Beauty/Fashion	3	76	50	40
Beauty/Fashion	2	62	48	40
Beauty/Fashion	1	52	34	28
Agriculture	3	2	2	2
Agriculture	2	2	2	2
Agriculture	1	4	2	2
Technology/Software	3	46	32	32
Technology/Software	2	28	18	18
Technology/Software	1	24	20	10

ROI Calculator for shark investors

Every shark wants to know in how many years their investment will be returned, so you must create a system for them, where shark will enter the name of startup and based on the total deal and equity given in how many years their principal amount will be returned and make their investment decision worth.

```
delimiter //
create procedure principal_returned (in startup varchar(100))
begin
    select Startup_Name,
        case
            when Accepted_Offer = 'No' or Accepted_Offer = 'No Offer Received.' then 'offer not accepted or not received'
            when Accepted_Offer = 'Yes' and (Yearly_Revenue_in_lakhs = '' or Yearly_Revenue_in_lakhs = 0) then 'previous data missing'
            else round(Total_Deal_Amount_in_lakhs/((Yearly_Revenue_in_lakhs*Total_Deal_Amount_in_lakhs)/100),2)
        end as 'result'
    from sharktank where Startup_Name = startup;
end
// delimiter ;

call principal_returned('BluePineFoods');
```

Startup_Name	result
BluePineFoods	1.05

Shark Investment Analysis

In the world of startup investing, we're curious to know which big-name investor, often referred to as "sharks," tends to put the most money into each deal on average. This comparison helps us see who's the most generous with their investments and how they measure up against their fellow investors.

```
select sharkname, round(avg(investment),2)  as 'average' from
(
SELECT `Namita_Investment_Amount_in lakhs` AS investment, 'Namita' AS sharkname FROM sharktank
WHERE `Namita_Investment_Amount_in lakhs` > 0
union all
SELECT `Vineeta_Investment_Amount_in_lakhs` AS investment, 'Vineeta' AS sharkname FROM sharktank
WHERE `Vineeta_Investment_Amount_in_lakhs` > 0
union all
SELECT `Anupam_Investment_Amount_in_lakhs` AS investment, 'Anupam' AS sharkname FROM sharktank
WHERE `Anupam_Investment_Amount_in_lakhs` > 0
union all
SELECT `Aman_Investment_Amount_in_lakhs` AS investment, 'Aman' AS sharkname FROM sharktank
WHERE `Aman_Investment_Amount_in_lakhs` > 0
union all
SELECT `Peyush_Investment_Amount_in_lakhs` AS investment, 'peyush' AS sharkname FROM sharktank
WHERE `Peyush_Investment_Amount_in_lakhs` > 0
union all
SELECT `Amit_Investment_Amount_in_lakhs` AS investment, 'Amit' AS sharkname FROM sharktank
WHERE `Amit_Investment_Amount_in_lakhs` > 0
union all
SELECT `Ashneer_Investment_Amount` AS investment, 'Ashneer' AS sharkname FROM sharktank
WHERE `Ashneer_Investment_Amount` > 0
)k group by sharkname;
```

sharkname	average
Namita	33.71
Vineeta	31.25
Anupam	29.99
Aman	34.89
peyush	35.92
Amit	35.31
Ashneer	25.67

Shark Investment Insights (stored procedure)

Develop a stored procedure that accepts inputs for the season number and the name of a shark. The procedure will then provide detailed insights into the total investment made by that specific shark across different industries during the specified season. Additionally, it will calculate the percentage of their investment in each sector relative to the total investment in that year, giving a comprehensive understanding of the shark's investment distribution and impact.

```
delimiter //
create procedure sharks_investment (in sea_num int, in sharkname varchar(10))
begin
    case
        when sharkname ='Namita' then
            set @total = (select sum(`Namita_Investment_Amount_in lakhs`) from sharktank where season_number = sea_num);
            select industry, sum(`Namita_Investment_Amount_in lakhs`) as 'amt_inv', round((100*sum(`Namita_Investment_Amount_in lakhs`)) / @total,2) as 'percentage_of_inv'
            from sharktank where season_number = sea_num group by industry;
        when sharkname ='Vineeta' then
            set @total = (select sum(`Vineeta_Investment_Amount_in_lakhs`) from sharktank where season_number = sea_num);
            select industry, sum(`Vineeta_Investment_Amount_in_lakhs`) as 'amt_inv', round((100*sum(`Vineeta_Investment_Amount_in_lakhs`)) / @total,2) as 'percentage_of_inv'
            from sharktank where season_number = sea_num group by industry;
        when sharkname ='Anupam' then
            set @total = (select sum(`Anupam_Investment_Amount_in_lakhs`) from sharktank where season_number = sea_num);
            select industry, sum(`Anupam_Investment_Amount_in_lakhs`) as 'amt_inv', round((100*sum(`Anupam_Investment_Amount_in_lakhs`)) / @total,2) as 'percentage_of_inv'
            from sharktank where season_number = sea_num group by industry;
    end case;
end
//
delimiter ;
call sharks_investment(2, 'Anupam');
```

industry	amt_inv	percentage_of_inv
Services	90	4.92
Food	480	26.23
Beauty/Fashion	520	28.42
Electronics	100	5.47
Entertainment	0	0
Vehicles/Electrical Vehicles	66.66	3.64