

Business Questions

Introduction

This file is a series of 16 questions that will allow you to answer business problems.

Each question will follow the following pattern:

- Business question
- SQL query (you must insert a screenshot)
- Answer (you must insert a screenshot)

The first two queries and answers have been completed for you. Run the queries yourself to confirm you get the same answer and add your screenshots below.

Section 1: Basic Summary Statistics

Question 1: Unique Countries

How many unique countries are there?

SQL Query

```
SELECT count(distinct CountryName)  
FROM PopStats;
```

Result

Unique Countries	
1	150

Question 2: Unique Cities

How many unique cities are there?

SQL Query

```
SELECT count(distinct City)
FROM FansPerCity
```

Result

	Unique Cities
1	46

Question 3: Unique Languages

How many unique languages are there?

SQL Query

```
SELECT count(distinct Language) FROM FansPerLanguage
```

Result(s)

	Unique_Languag
1	41

Question 4: Global Page Reach

What is the daily average reach of the posts (i.e. DailyPostsReach) on the global page over the period?

SQL Query

```
SELECT ROUND(AVG(DailyPostsReached),2) AS Daily_Average_Reach FROM GlobalPage
```

Result(s)

Daily_Average_Reach
1862816.03

Question 5: Global Page Likes

What is the daily average engagement rate (i.e. NewLikes) on the global page over the period?
Round the result to 2 decimal places.

SQL Query

```
SELECT ROUND(AVG(NewLikes),2) AS New_Likes FROM GlobalPage
```

Result(s)

New_Likes
8942.56

Section 2: Location Analysis

In these queries, make sure you select the latest date in the main table. E.g:

```
SELECT max(date) from FansPerCountry
```

Then use the result of this in the WHERE clause of your main query.

Question 6: Top 10 countries by Number of Fans

What are the top 10 countries (considering the number of fans)? Show a table of results containing the following columns:

- CountryCode
- Country Name
- NumberOfFans

SQL Query

```
SELECT CountryCode, CountryName, NumberOfFans
FROM FansPerCountry
INNER JOIN Popstats
ON FansPerCountry.CountryCode = PopStats.CountryCode
WHERE Date = (SELECT max(date) FROM FansPerCountry)
ORDER BY NumberOfFans DESC
LIMIT 10
```

Result(s)

	CountryCode	CountryName	NumberOfFans
1	CI	Ivory Coast	112160
2	CM	Cameroon	102211
3	SN	Senegal	83561
4	FR	France	73252
5	MG	Madagascar	72956
6	CD	Democratic Republic of the Congo	50705
7	BF	Burkina Faso	43500
8	ML	Mali	40578
9	DZ	Algeria	39093
10	GN	Guinea	36821

Question 7: Top 10 countries by Penetration Ratio

What are the top 10 countries by penetration ratio (i.e. the % of the country population that are fans)? Show a table of results containing the following columns:

- CountryName
- PenetrationRatio
- NumberOfFans
- Population

SQL Query

```
SELECT CountryName,  
NumberOfFans,  
Population,  
ROUND(((1.0*NumberOfFans)/(1.0*Population))*100,7) AS Pen_Ratio  
FROM FansPerCountry INNER JOIN Popstats ON FansPerCountry.CountryCode =  
PopStats.CountryCode  
WHERE date = (SELECT max(date) FROM FansPerCountry)  
ORDER BY NumberOfFans DESC  
LIMIT 10
```

Result(s)

	CountryName	NumberOfFan:	Population	Pen Ratio
1	Reunion	20885	866506	2.4102545
2	French Polynesia	5148	283007	1.8190363
3	New Caledonia	5032	280460	1.7941953
4	Mauritius	24210	1364283	1.7745585
5	Martinique	5427	376480	1.4415108
6	Guadeloupe	5379	395700	1.3593632
7	Gabon	23954	2119036	1.1304197
8	Mayotte	1983	270372	0.7334339
9	Comoros	4925	821164	0.5997584
10	French Guiana	1687	296711	0.5685667

Question 8: Bottom 10 Cities by Number of Fans

What are the bottom 10 cities (considering the number of fans) among countries with a population over 20 million?

This could be considered our growth potential.

Show a table of results containing the following columns:

- CountryName
- City
- NumberOfFans
- Population

SQL Query

```
SELECT
CountryName,
City,
NumberOfFans,
Population
FROM FansPerCity INNER JOIN Popstats ON FansPerCity.CountryCode =
PopStats.CountryCode
WHERE (Population >= 20000000) AND date = (SELECT max(date) FROM FansPerCity)
ORDER BY NumberOfFans ASC
LIMIT 10
```

Result(s)

	CountryName	City	NumberOfFans	Population
1	Algeria	Bejaia	2301	41657488
2	Cameroon	Ngaoundere	2318	25640965
3	Madagascar	Fianarantsoa	2366	25683610
4	Algeria	Tizi Ouzou	2524	41657488
5	Canada	Montreal	2887	35881659
6	Algeria	Oran	2920	41657488
7	Ivory Coast	Bouake	3376	24290000
8	Ivory Coast	Cocody	3647	24290000
9	Morocco	Casablanca	3951	34314130
10	Angola	Luanda	4614	30355880

Section 3: Fan Analysis

Question 9: Analysis by age group (split of fans)

What is the split of page fans across age groups (in %)?

Show a table of results containing the following columns:

- AgeGroup
- PercentageOfFans

SQL Query

```
SELECT AgeGroup,
(1.0 * SUM(NumberOfFans) ) --Numerator is number of fans in age group
/
(1.0 * (SELECT SUM(NumberOfFans) FROM FansPerGenderAge) ) * 100 --denom is total
fans * 100
AS PercentageOfFans
FROM FansPerGenderAge
GROUP BY AgeGroup;
```

Result(s)

AgeGroup	PercentageOfFans
13-17	2.13197938169817
18-24	21.30039293198776
25-34	35.73366398752555
35-44	19.39725957600248
45-54	9.47496562259656
55-64	5.02402272752625
65+	6.93771577266323

Question 10: Analysis by gender (split of fans)

What is the split of page fans by gender (in %)?

Show a table of results containing the following columns:

- Gender
- PercentageOfFans

SQL Query

```
SELECT Gender,
(1.0 * SUM(NumberOfFans) ) --Numerator is number of fans in age group
/
```

```
(1.0 * (SELECT SUM(NumberOfFans) --denom is total fans
FROM FansPerGenderAge) )
* 100 --As a Percentage
AS PercentageOfFans
FROM FansPerGenderAge
GROUP BY Gender;
```

Result(s)

Gender	PercentageOfFans
F	56.46119943597241
M	43.44411007522478
U	0.09469048880281

Section 4: Language Analysis

Question 11: English Speaking Fans

What is the number of the fans that have declared English as their primary language ?

SQL Query

```
SELECT
language,
SUM(NumberOfFans) AS NumberOfFans
FROM FansPerLanguage
WHERE Language = "en"
GROUP BY Language
```

Result(s)

	language	NumberOfFans
1	en	1347752

Question 12: English Speaking Fans Percentage

What is the percentage of the fans that have declared English as their primary language?

SQL Query

```
SELECT
(1.0*SUM(NumberOfFans))
/
(1.0*(SELECT SUM(NumberOfFans) FROM FansPerLanguage))
*100 AS NumberOfFans

FROM FansPerLanguage
WHERE Language = "en"
GROUP BY Language
```

Result(s)

	NumberOfFans
1	5.08050979857463

Question 13: Buying Power of English Speakers

Based on the number of fans who have declared English as their primary language and living in the US, what is the potential buying power that can be accessed? (Please use the average income data per country for this question. It is estimated that on average, 0.01% of the annual income is dedicated to online magazine subscriptions in the US).

SQL Query

```
SELECT
AverageIncome*.0001 *
FansPerCountry.NumberOfFans -- (SELECT NumberOfFans FROM FansPerLanguage
WHERE language != "en" AND CountryCode = "US")
AS BuyingPower
FROM PopStats JOIN FansPerCountry ON PopStats.CountryCode =
FansPerCountry.CountryCode,
FansPerLanguage ON PopStats.CountryCode = FansPerLanguage.CountryCode
WHERE CountryName = "United states" AND Language = "en"
GROUP BY Language
```

Result(s)

BuyingPower
55675.814116

Section 5: Fan Engagement

Question 14: Engagement per day of the week

What is the split of the EngagedFans across the days of the week (monday, tuesday,...)?

Give the result as a table with the following columns:

- DayOfWeek
- PercentageSplit

Based on the results, what is the best day of the week to publish posts?

SQL Query

```
SELECT
((1.0*SUM(EngagedFans)) / (1.0*(SELECT SUM(EngagedFans) FROM PostInsights)))
* 100.0 AS PercentageSplit,
CASE
WHEN strftime('%w', CreatedTime) = '0' THEN "Sunday"
WHEN strftime('%w', CreatedTime) = '1' THEN "Monday"
WHEN strftime('%w', CreatedTime) = '2' THEN "Tuesday"
WHEN strftime('%w', CreatedTime) = '3' THEN "Wednesday"
WHEN strftime('%w', CreatedTime) = '4' THEN "Thursday"
WHEN strftime('%w', CreatedTime) = '5' THEN "Friday"
WHEN strftime('%w', CreatedTime) = '6' THEN "Saturday"
END AS DaysOfWeek
FROM PostInsights
GROUP BY DaysOfWeek
ORDER BY strftime('%w', CreatedTime)
```

Result(s)

	PercentageSplit	DaysOfWeek
1	12.08265316495928	Sunday
2	19.2340320614915	Monday
3	18.66981851919799	Tuesday
4	15.38064695551377	Wednesday
5	6.32203114538212	Thursday
6	8.58519525165548	Friday
7	19.72562290179985	Saturday

Based on the results, the best day of the week to post is Saturday

Question 15: Engagement per time of day

What is the split of the EngagedFans by time of the day? Split the day into the following time ranges (call the column TimeRange):

- 05:00 - 08:59
- 09:00 - 11:59
- 12:00 - 14:59

- 15:00 - 18:59
- 19:00 - 21:59
- 22:00 or later

Give the result as a table with the following columns:

- TimeRange
- Percentage

Based on the results, what is the best time of the day to publish posts?

SQL Query

```
SELECT
(1.0*SUM(EngagedFans)) / (1.0*(SELECT SUM(EngagedFans) FROM PostInsights))
* 100.0 AS PercentageSplit,
CASE
WHEN 1*strfime('%H', CreatedTime) >= 5 AND 1*strfime('%H', CreatedTime) < 9 THEN
'5AMTo9AM'
WHEN 1*strfime('%H', CreatedTime) >= 9 AND 1*strfime('%H', CreatedTime) < 12
THEN '9AMTo12PM'
WHEN 1*strfime('%H', CreatedTime) >= 12 AND 1*strfime('%H', CreatedTime) < 15
THEN '12PMTto3PM'
WHEN 1*strfime('%H', CreatedTime) >= 15 AND 1*strfime('%H', CreatedTime) < 19
THEN '3PMTto7PM'
WHEN 1*strfime('%H', CreatedTime) >= 19 AND 1*strfime('%H', CreatedTime) < 22
THEN '7PMto10PM'
--WHEN 1*strfime('%H', CreatedTime) >= 22 OR 1*strfime('%H', CreatedTime) < 5
THEN '10PMTto5AM'
ELSE '10PMTto5AM'
END AS TimeRange
FROM PostInsights
GROUP BY TimeRange
Order By CreatedTime
```

Result(s)

PercentageSplit	TimeRange
33.96124997516922	5AMTo9AM
15.33583471704022	9AMTo12PM
17.40415030270171	3PMTo7PM
13.12151419794273	7PMto10PM
7.47879452157239	10PMTto5AM
12.69845628557373	12PMTto3PM

Based on the results, the best time of time of day to post is between 5 am to 9 am