Analyze Data

Sayali Kedar Oak

```
library(RMySQL)
```

```
## Loading required package: DBI
```

```
dbcon = dbConnect(RMySQL::MySQL(),host='localhost',port=3306,user='root',
                  password='root123')
```

```
USE practicum2DB
```

Report Title : Popular Journals for a given time period

Analytical Query I:

Top five journals with the most articles published in them for the time period. Think about how you would best present this information so it is easily usable by an analyst or manager.

Consider the time period as 1975 - 1980 The following query gives information about the top 5 journals in which most articles were published during this time period. The best way to present this information will be in the form of a table like the one from the output of the query, or a bar graph.

The following report provides information about the most popular journals for authors. Since it gives information about the journals which published the most number of articles for a given time period.

```
SELECT JournalTitle, SUM(JournalFacts.NumArticlesPerYear) as TotalArticles
FROM JournalFacts
JOIN JournalDimension ON JournalFacts.JournalID = JournalDimension.JournalID
JOIN TimeDimension ON JournalFacts.TimeID = TimeDimension.TimeID
WHERE TimeDimension.Year BETWEEN 1975 AND 1980
GROUP BY JournalDimension.JournalTitle
ORDER BY TotalArticles DESC
LIMIT 5;
```

Table 1: 5 records

JournalTitle	TotalArticles
Biochemical pharmacology	188
Biochemical and biophysical research communications	188
Biochemical medicine	94
Arzneimittel-Forschung	94
British journal of anaesthesia	94

Report title: Articles published in journals per quarter.

##Analytical Query II: Number of articles per journal per year broken down by quarter. Again, think of a good way to show this. Is a table better than a visualization or is it better to use a visualization.

In this case, a visualization is a better way to present the data provided. This is because they table may cause confusion since there is a lot of numerical data present.

The following report gives information about the number of articles published in journals per quarter.

```
SELECT JournalTitle, TimeDimension.Year, TimeDimension.Quarter,
SUM(NumArticlesPerQuarter) as TotalArticles
FROM JournalFacts
JOIN JournalDimension ON JournalFacts.JournalID = JournalDimension.JournalID
JOIN TimeDimension ON JournalFacts.TimeID = TimeDimension.TimeID
GROUP BY JournalTitle, TimeDimension.Year, TimeDimension.Quarter
ORDER BY JournalTitle AND TimeDimension.Quarter AND TimeDimension.Year;
```

Table 2: Displaying records 1 - 10

JournalTitle	Year	Quarter	TotalArticles
Biochemical medicine	1975	2	13
Biochemical and biophysical research communications	1975	3	106
Biochemical pharmacology	1975	3	106
Arzneimittel-Forschung	1975	3	53
British journal of anaesthesia	1975	3	53
The British journal of clinical practice	1975	2	13
The British journal of dermatology	1975	2	13
British journal of haematology	1975	1	6
The British journal of nutrition	1975	3	53
Biochemistry	1975	3	53

```
s <- dbDisconnect(dbcon)
```