MGMT 582 - Project Section 2 – Group 3

EVENT MANAGEMENT BUSINESS DATABASE SOLUTION

Submitted by: -

1. Esha Chouthe
2. Rohit Soans
3. Padma Dwivedi
4. Devansh Batra
5. Kshitij Virdi
6. Adithya Bharadwaj Umamahesh

1.1 Client Overview and Background

Lafayette - West Lafayette is home to an endless array of fun events all year long. Its calendar is full of live entertainment, social events, festivals, art, and cultural events, and more.

Our client is an **Event Manager** who has been hosting events in West Lafayette and downtown Lafayette for a year now. He currently tracks data on excel and has no systematic and sophisticated database structure. As he is solely responsible for end-to-end management of all events including procurement of suppliers and materials, vendor selection and choosing locations for the events, he is unable to efficiently track all his expenses given such steep deadlines.

Our objective, as a team, is to create visibility for his **profitability** and to **optimize** his **business** to **increase profitability**. Besides this, we also have investigated **seasonal trends**, **supplier data** and **customer demography** that could help optimize the business.

The company hosts events like parties, conferences, concerts, and various team events for corporations. Customer base includes both individual parties and organizations. A few of the primary locations for events hosted by the company are Columbia Park, Wea Creek Orchard, Happy Hollows, etc. Apart from the outdoors, a few events are held at Home/Office locations as well. For organizing events, the company hires some staff members who are both permanent and contractual. The company further buys their supplies from various vendors who specialize in distribution of various products and services and charge him for the same on an event basis. Majority of these items are event specific and can be customized to host an unforgettable event for the hosts. For example, personalized return gift for a kid's birthday based on Marvel characters. To promote a few crowd-based events, the company runs commercial advertisements on multiple channels, including social media platforms. The advertisements are add-on charges to hosts. The company curates the content as well as channel to promote the advertisement to increase exposure for the event. Advertisement can increase the engagement to the event by increasing attendance, post event surveys, etc. Company is committed to convert the hosts' dream events into reality. Their aim is to maximize reach by increasing the number of new customers and maximize the frequency by increasing the number of events with an entire customer base.

1.2 Dataset Description

Raw Table was obtained from the client which included the following columns in 1 – NF form:

EventID, Event_Name, Type, location, charge_to_cust, date, CustID, OverallRating, Cust_name, Age, AdID, Medium, Cost, CustPhone, CustBudget, Vendor_ID, Item, unit_price ($), qty, lag_time, StaffID, Hours, Staff_First_Name, Staff_Last_Name, Age, Type, Position, Department, Bill_rate_per_hour.

1.2.1 Business Case and Cardinalities

- Each event is hosted by one and only one customer.
- A customer can have multiple events of diverse types.
- A customer can be an individual or an organization or both (primary contact of organization has also hired the company to host personal events).
- One or multiple vendors can provide items for one event and each event can have one or more vendor's providing items.
- An event is marketed by multiple advertisements mediums whereas one advertisement is done for one and only one event. A few events do not require any advertisements.
- Each event will be allocated to one or more staff
- Each member of staff could either be an employee or be hired on a contractual basis for large events.
- Each member of staff could be placed for one or more. Few contractual and employed staff are not allocated to events based on their availability.

1.3 Problem Statement

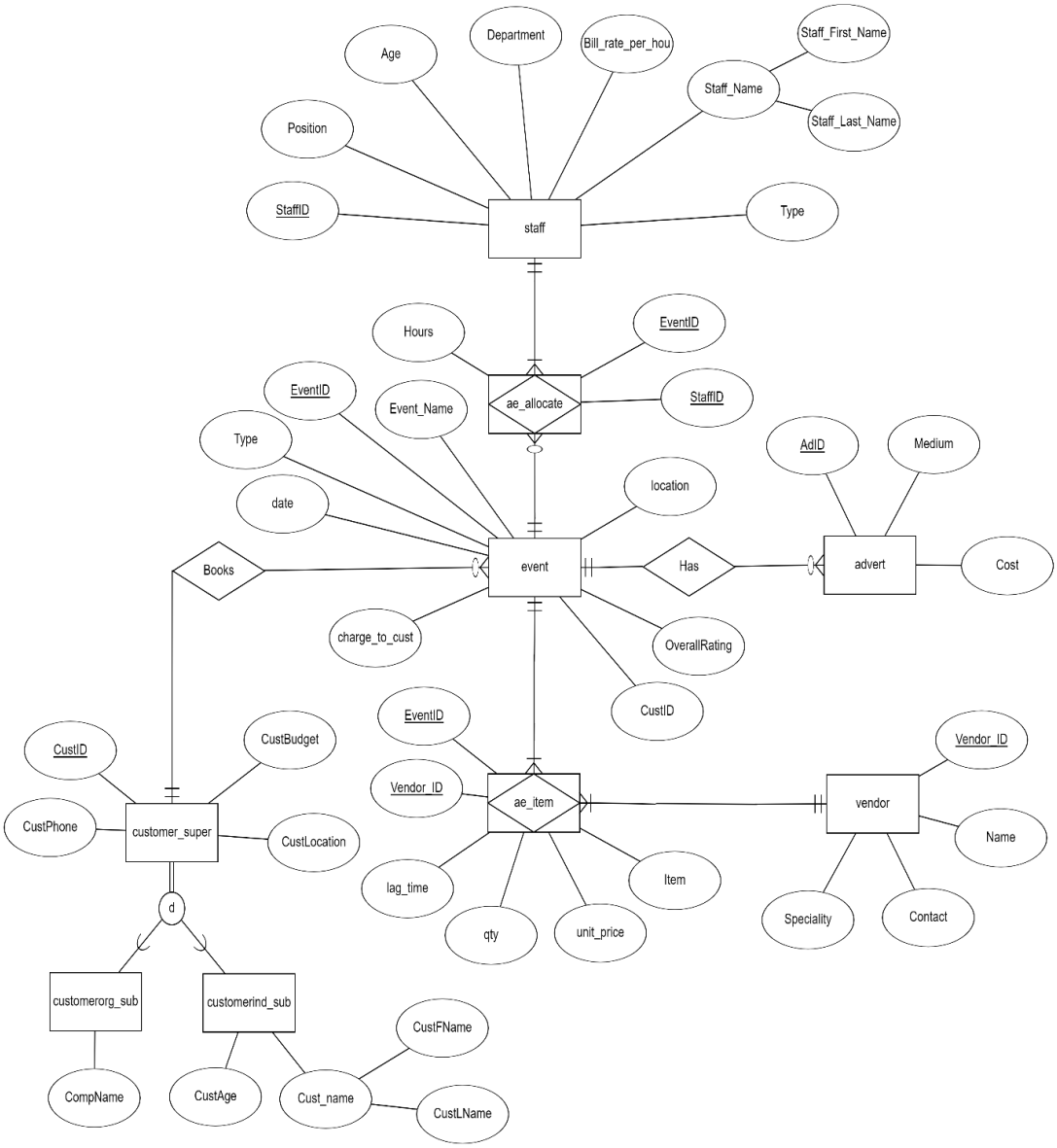
Client is unable to efficiently track all his expenses and events including procurement of materials, vendor selection and choosing locations for the events. This adversely affects his profitability.

1.4 Project Objectives

Our objective is not only to create visibility on his profitability and improve it, but also to efficiently manage his Database so he can track all the series of events and deduce inferences.

In addition to this, we will look for seasonal trends, supplier data and customer demography that could help optimize the business.

1.6 Conceptual Data Modelling: ERD



1.7 Relational Schema

customerorg_sub(CustID, CompName) customerind_sub(CustID, CustFName, CustLName, Age)

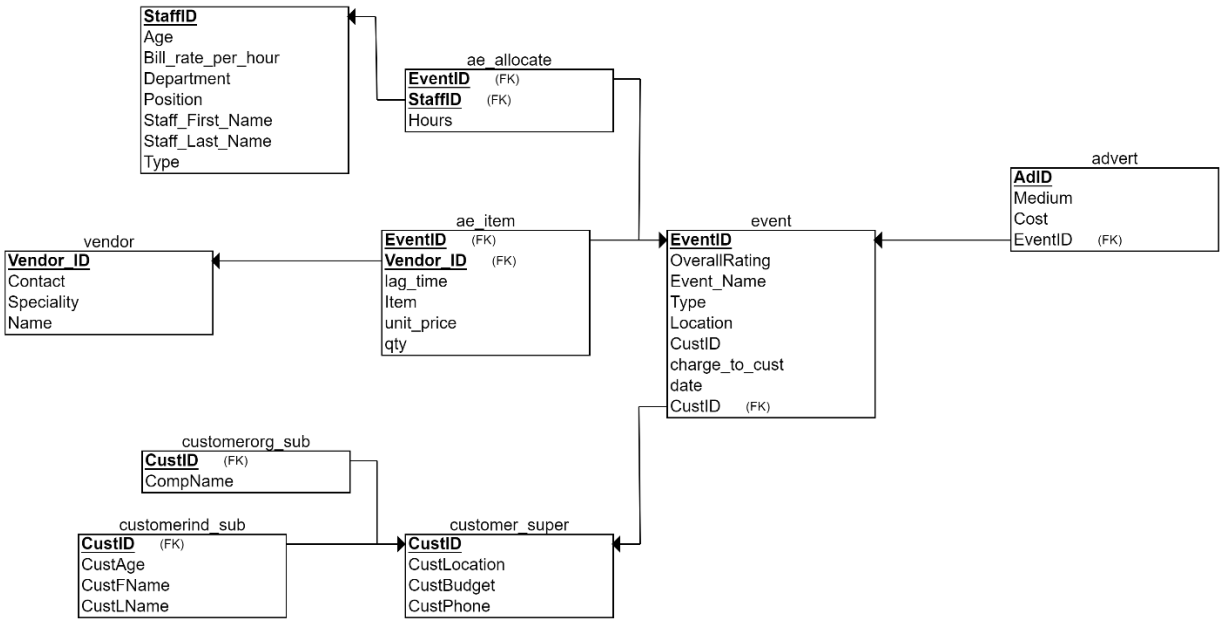
customer_super(CustID, CustPhone, CustLocation, CustBudget)

event(EventID, Event_Name, Type, location, charge_to_cust, date, OverallRating, CustID)

advert(AdID, Medium, Cost, EventID) ae_item(EventID, Vendor_ID, lag_time, Item, qty, unit_price)

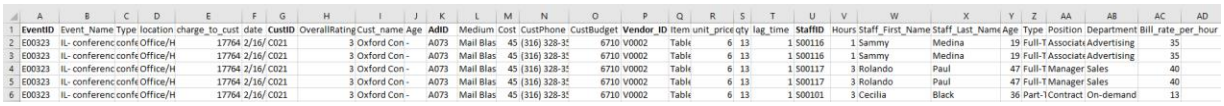
vendor(Vendor_ID, Contact, Name, Specialty) ae_allocate(StaffID, EventID, hour)

staff(StaffID, Staff_First_Name, Staff_Last_Name, Age, Position, Department, Type, Bill_rate_per_hour)



1.8 Data Model and Design Choices:

The data that the client provided is as below:



EventID	Event_Name	Type	location	charge_to_cust	date	CustID	OverallRating	Cust_name	Age	AdID	Medium	Cost	CustPhone	CustBudget	Vendor_ID	Item	unit_price	qty	lag_time	StaffID	Hours	Staff_First_Name	Staff_Last_Name	Age	Type	Position	Department	Bill_rate_per_hour
E00323	IL- conferenc	confe	Office/H	17764	2/16/	C021	3	Oxford Con -		A073	Mail Blas	45	(316) 328-3	6710	V0002	Table	6	13	1	S00116	1	Sammy	Medina	19	Full-T	Associate	Advertising	35
E00323	IL- conferenc	confe	Office/H	17764	2/16/	C021	3	Oxford Con -		A073	Mail Blas	45	(316) 328-3	6710	V0002	Table	6	13	1	S00116	1	Sammy	Medina	19	Full-T	Associate	Advertising	35
E00323	IL- conferenc	confe	Office/H	17764	2/16/	C021	3	Oxford Con -		A073	Mail Blas	45	(316) 328-3	6710	V0002	Table	6	13	1	S00117	3	Rolando	Paul	47	Full-T	Manager	Sales	40
E00323	IL- conferenc	confe	Office/H	17764	2/16/	C021	3	Oxford Con -		A073	Mail Blas	45	(316) 328-3	6710	V0002	Table	6	13	1	S00117	3	Rolando	Paul	47	Full-T	Manager	Sales	40
E00323	IL- conferenc	confe	Office/H	17764	2/16/	C021	3	Oxford Con -		A073	Mail Blas	45	(316) 328-3	6710	V0002	Table	6	13	1	S00101	3	Cecilia	Black	36	Part-1	Contract	On-demand	13

Vendor_ID	Name	Contact	Specialty
V0001	Duncan Enterprises	7657755100	Birthdays
V0002	Arzoo & Bros	7657755330	Concerts

Table 1: (EventID, Event_Name, Type, location, charge_to_cust, date, CustID, OverallRating, Cust_name, Age, AdID, Medium, Cost, CustPhone, CustBudget, Vendor_ID, Item, unit_price, qty, lag_time, StaffID, Hours, Staff_First_Name, Staff_Last_Name, Age, Type, Position, Department, Bill_rate_per_hour)

Table 2: (Vendor_ID, Name, Contact, Specialty)

We can see that the table contains all the atomic and unique values and hence we concluded that the table client provided is in 1NF.

We found the partial dependencies in the dataset, to convert this into 2NF, which are as below,

Partial Dependencies:

EventID, CustID → {Event_Name, Type, location, charge_to_cust, date, OverallRating}

CustID → {CustName, Age, location, Phone, Budget}

AdID, EventID → {Medium, Cost}

Vendor_ID → {Name, Contact, Specialty}

EventID, Vendor_ID → {Item, unit_price, qty, lag_time}

StaffID → {Staff_First_Name, Staff_Last_Name, Age, Position, Department, Type, Bill_rate_per_hour}

StaffID, EventID → {hour}

2NF form ---------->

2NF Form:

event(EventID, Event_Name, Type, location, charge_to_cust, date, OverallRating, CustID)

customer(CustID, CustName, Age, location, Phone, Budget)

advert(AdID, Medium, Cost, EventID)

vendor(Vendor_ID, Contact, Name, Specialty)

ae_item(EventID, Vendor_ID, lag_time, Item, unit_price, qty)

staff(StaffID, Staff_First_Name, Staff_Last_Name, Age, Position, Department, Type, Bill_rate_per_hour)

ae_allocate(StaffID, EventID, hour)

We can see that the Customer table has two categories: Individual and Organizations. The Organizations has only Company Name in their ‘Name’ attribute whereas the Individuals have First and Last name in the ‘Name’ attribute along with an ‘Age’ attribute. Hence, we decided to split the customer entity into 2 sub types: Individual and Organization.

There are Transitive dependencies in the data that correspond to the super-subtype entities and hence the transitive dependencies along with data in 3NF are as follows:

Transitive dependencies:

CustomerID, CustName → (Age) We removed Transitive dependency to convert it to 3NF.

3NF Form:

event(EventID, Event_Name, Type, location, charge_to_cust, date, OverallRating, CustID)

customer_super(CustID, CustLocation, CustPhone, CustBudget)

customerind_sub(CustID, CustFName, CustLName, Age)

customerorg_sub(CustID, CompName)

advert(AdID, Medium, Cost, EventID)

vendor(Contact, Name, Specialty)

ae_item(EventID, Vendor_ID, lag_time, Item, unit_price, qty)

staff(StaffID, Staff_First_Name, Staff_Last_Name, Age, Position, Department, Type, Bill_rate_per_hour)

ae_allocate(StaffID, EventID, hour)

This becomes our final data model which we will use for our project hereon

1.9 Queries and Description (Analysis)

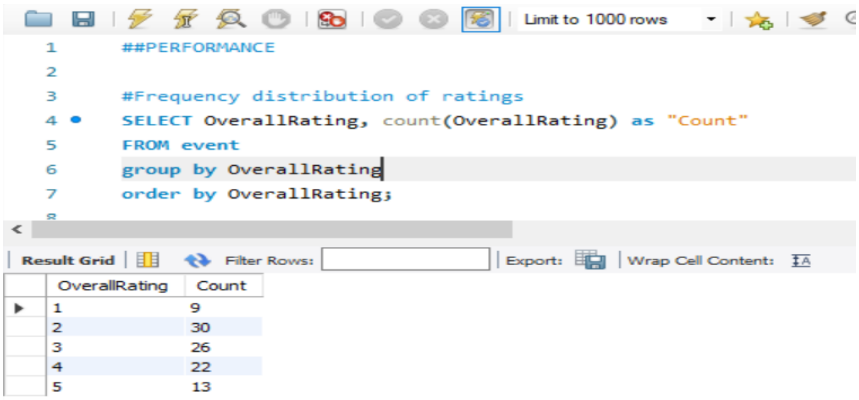
We have analyzed various trends **which affect overall profitability** and queried data to deduce inferences for the following parameters:

- **Performance**
- **Advertising Spending**
- **Seasonality**
- **Marketing Spend**
- **Customer Analysis**
- **Vendor Analysis**

1.9.1 Performance

Query 1

Business Case: What was the distribution of the performance rating between 1-5(1=lowest, 5=highest)



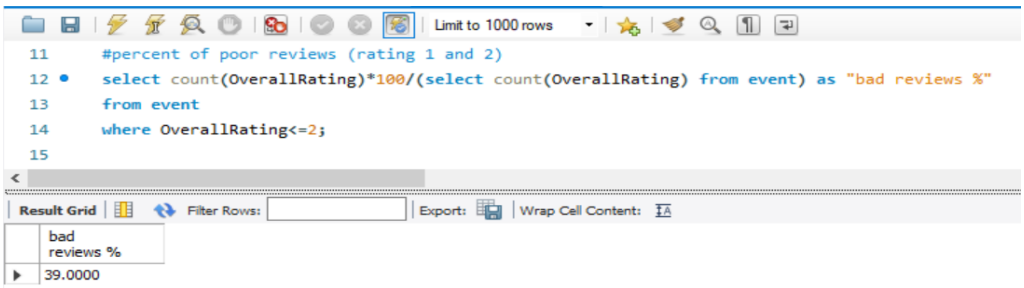
```
##PERFORMANCE

#Frequency distribution of ratings
SELECT OverallRating, count(OverallRating) as "Count"
FROM event
group by OverallRating
order by OverallRating;
```

OverallRating	Count
1	9
2	30
3	26
4	22
5	13

Inference: From this we can see that most of the ratings range from 2 to 4. Hence there is room for improvement.

Query 2

Business Case: How many customers left poor ratings?



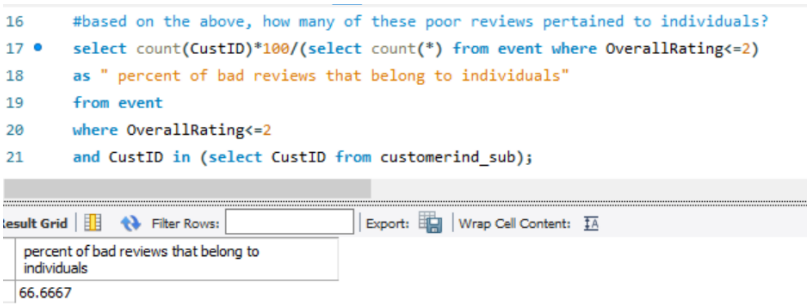
```
#percent of poor reviews (rating 1 and 2)
select count(OverallRating)*100/(select count(OverallRating) from event) as "bad reviews %"
from event
where OverallRating<=2;
```

bad reviews %
39.0000

Inference: If we categorize a 1 and 2 rating as “Poor”, we see that 39 % of the client’s ratings are poor

Query 3

Business Case: Among those customers who left poor ratings, how many of them were given by individuals as opposed to organizations?



```
#based on the above, how many of these poor reviews pertained to individuals?
select count(CustID)*100/(select count(*) from event where OverallRating<=2)
as " percent of bad reviews that belong to individuals"
from event
where OverallRating<=2
and CustID in (select CustID from customerind_sub);
```

percent of bad reviews that belong to individuals
66.6667

Inference: When we further investigate the poor ratings, we notice that two-thirds of poor ratings were given by individual customers as opposed to institutional customers. This means Clay must focus more on the individual customers' experience.

1.9.2 Advertising Spending

Query 4

Business Case: How was the advertising budget appropriated among various mediums?

```
#on which medium was the most amount of money($) spent?
select Medium, sum(Cost) as 'Total Cost' from advert
group by Medium
order by sum(Cost) desc;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Medium	Total Cost
Radio	2106
Social Media	887
Newspaper	780
Mail Blast	443

Inference: Our client spent the most money (USD) advertising over the radio.

Query 5

Business Case: how much sales were generated from each form of advertising?

```
#how did each form of advertising media affect the sales?
select a.Medium, SUM(e.charge_to_cust) as "Event Sales"
from advert a
join event e
on a.EventID=e.EventID
group by a.Medium
order by SUM(e.charge_to_cust) desc ;
```

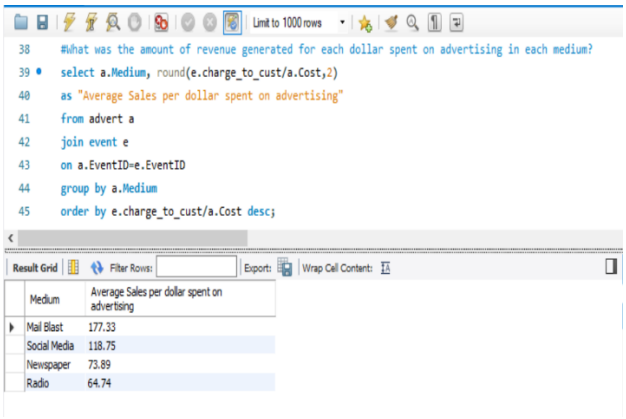
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Medium	Event Sales
Radio	205113
Social Media	154792
Mail Blast	150058
Newspaper	78465

Inference: Radio advertising also generated the most sales (in USD)

Query 6

Business Case: However, it would be more worthwhile to look at the average sales generated from each dollar spent on advertising from each form of media.



```
38  #what was the amount of revenue generated for each dollar spent on advertising in each medium?
39  select a.Medium, round(e.charge_to_cust/a.Cost,2)
40  as "Average Sales per dollar spent on advertising"
41  from advert a
42  join event e
43  on a.EventID=e.EventID
44  group by a.Medium
45  order by e.charge_to_cust/a.Cost desc;
```

Medium	Average Sales per dollar spent on advertising
Mail Blast	177.33
Social Media	118.75
Newspaper	73.89
Radio	64.74

Inference: It is surprising that despite having spent most of the ad budget on radio advertising, it has the lowest revenue per dollar of advertising expended. **Hence Clay would have to reduce the expenditure on the radio and increase the same on Mail Blast to Maximize Profits.**

Query 7

Business Case: what was the event-wise revenue generated per dollar spent on advertising?

```
48  #what was the event-wise revenue generated per dollar spent on advertising?
49  select e.Type, round(e.charge_to_cust/a.Cost,2)
50  as "Event-wise average Sales per dollar spent on advertising"
51  from advert a
52  join event e
53  on a.EventID=e.EventID
54  group by e.Type
55  order by e.charge_to_cust/a.Cost desc;
```

Type	Event-wise average Sales per dollar spent on advertising
conference	296.47
birthday	115.20
party	98.11
concerts	74.63
team event	73.89
others	64.74

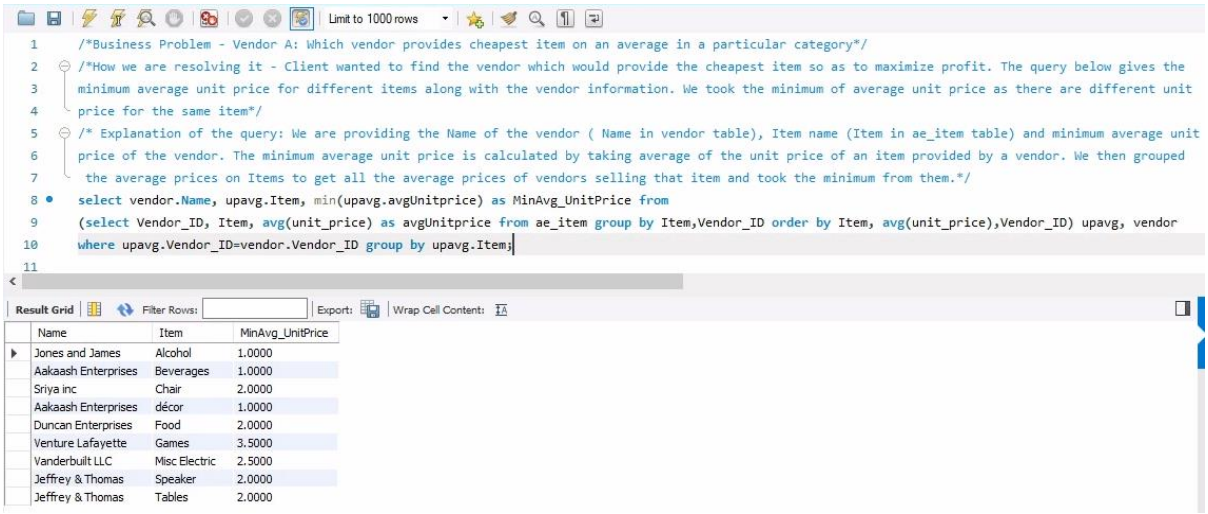
Inference: Clay is getting a better return on advertising expenses by planning conferences.

Hence, he should target more conference events to minimize a potential opportunity loss.

1.9.3 Vendor Analysis:

Query 8

Business Case: Which vendor provides the cheapest item on average in a particular category?



```
/*Business Problem - Vendor A: Which vendor provides cheapest item on an average in a particular category*/
/*How we are resolving it - Client wanted to find the vendor which would provide the cheapest item so as to maximize profit. The query below gives the
minimum average unit price for different items along with the vendor information. We took the minimum of average unit price as there are different unit
price for the same item*/
/* Explanation of the query: We are providing the Name of the vendor ( Name in vendor table), Item name (Item in ae_item table) and minimum average unit
price of the vendor. The minimum average unit price is calculated by taking average of the unit price of an item provided by a vendor. We then grouped
 the average prices on Items to get all the average prices of vendors selling that item and took the minimum from them.*/
select vendor.Name, upavg.Item, min(upavg.avgUnitprice) as MinAvg_UnitPrice from
(select Vendor_ID, Item, avg(unit_price) as avgUnitprice from ae_item group by Item,Vendor_ID order by Item, avg(unit_price),Vendor_ID) upavg, vendor
where upavg.Vendor_ID=vendor.Vendor_ID group by upavg.Item;
```

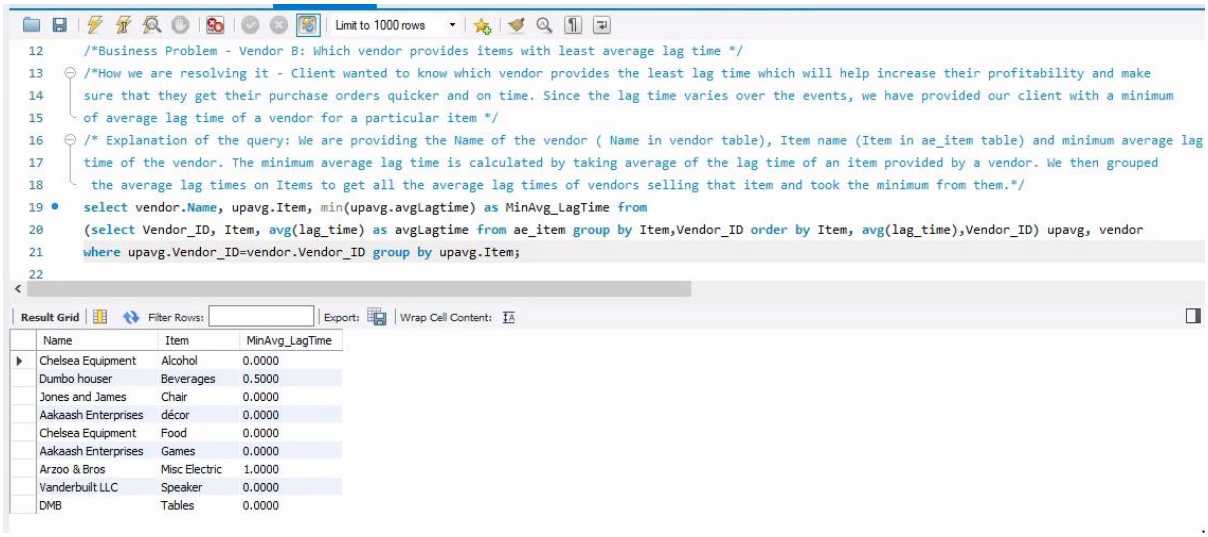
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Name	Item	MinAvg_UnitPrice
Jones and James	Alcohol	1.0000
Aakaash Enterprises	Beverages	1.0000
Sriya inc	Chair	2.0000
Aakaash Enterprises	décor	1.0000
Duncan Enterprises	Food	2.0000
Venture Lafayette	Games	3.5000
Vanderbuilt LLC	Misc Electric	2.5000
Jeffrey & Thomas	Speaker	2.0000
Jeffrey & Thomas	Tables	2.0000

Inference: We see that there are a lot of vendors providing the items at relatively cheaper prices and this report helps our client to compare the future vendor costs as well. This will help increase profitability by minimizing our spend on items/equipments.

Query 9

Business Case: Which vendor provides items with least average lag time?



```
/*Business Problem - Vendor B: Which vendor provides items with least average lag time */
/*How we are resolving it - Client wanted to know which vendor provides the least lag time which will help increase their profitability and make
sure that they get their purchase orders quicker and on time. Since the lag time varies over the events, we have provided our client with a minimum
of average lag time of a vendor for a particular item */
/* Explanation of the query: We are providing the Name of the vendor ( Name in vendor table), Item name (Item in ae_item table) and minimum average lag
time of the vendor. The minimum average lag time is calculated by taking average of the lag time of an item provided by a vendor. We then grouped
 the average lag times on Items to get all the average lag times of vendors selling that item and took the minimum from them.*/
select vendor.Name, upavg.Item, min(upavg.avgLagtime) as MinAvg_LagTime from
(select Vendor_ID, Item, avg(lag_time) as avgLagtime from ae_item group by Item,Vendor_ID order by Item, avg(lag_time),Vendor_ID) upavg, vendor
where upavg.Vendor_ID=vendor.Vendor_ID group by upavg.Item;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Name	Item	MinAvg_LagTime
Chelsea Equipment	Alcohol	0.0000
Dumbo houser	Beverages	0.5000
Jones and James	Chair	0.0000
Aakaash Enterprises	décor	0.0000
Chelsea Equipment	Food	0.0000
Aakaash Enterprises	Games	0.0000
Arzoo & Bros	Misc Electric	1.0000
Vanderbuilt LLC	Speaker	0.0000
DMB	Tables	0.0000

Inference: With this report, our client will be able to make the right choice of selecting a vendor as a less lag time will only ensure on time procurement and boost efficiency and profitability.

1.9.4 Profitability:

Query 10

Business Case: What is the Profit for each event?

```
/*Business Problem - Profitability: What is the Profit for each event */
/*How we are resolving it - Client was unable to view his profit against each event. We provided this functionality using the query below which now gives our
client visibility of profit across all events. We created this report by displaying the Event Details, Customer ID, Revenue, Cost and Profit details*/
/*Explanation of the query: The Profit for each event can be calculated by the Revenue earned from that event and later deducting the costs associated
with the event. The Revenue earned is the amount client charged the customer, that is stored in our Event table in the column charge_to_cust. The cost of
each event is a total of cost of all items purchased, cost of staff allocated for the event and the cost of advertisements for the event.
The cost of advertisements is stored in the advert table under the column cost(against event id). The cost of staff can be calculated by product of the
number of hours a staff worked for an event (Hours in ae_allocate table) and the amount a staff charges per hour (Bill_rate_per_hour in staff table).
The cost of item is calculated as product of the unit price of an item (unit_price in ae_item table) and quantity ordered (qty in ae_item table). These are
all calculated against the event id and the summed together to get the Total cost.  */
```

```
SELECT Revenue.*, Cost.Cost , (Revenue - Cost) AS Profit FROM
(select EventID, CustID, Event_Name, charge_to_cust AS Revenue from event) AS Revenue
inner join
(select ID, SUM(Cost) AS Cost
FROM(
select ae_item.EventID AS ID, sum(ae_item.unit_price*ae_item.qty) as Cost from ae_item group by ae_item.EventID
UNION ALL
select advert.EventID AS ID, sum(advert.Cost) as Cost from advert group by advert.EventID
UNION ALL
select ae_allocate.EventID AS ID, sum(ae_allocate.Hours*staff.Bill_rate_per_hour) as Cost
from ae_allocate JOIN staff ON ae_allocate.StaffID=staff.StaffID group by ae_allocate.EventID) AS A
group by ID) AS Cost
on Revenue.EventID = Cost.ID;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

EventID	CustID	Event_Name	Revenue	Cost	Profit
E00302	C026	KC- team event Jan-26	13300	2584	10716
E00304	C014	OQ- others Jan-28	17481	3068	14413
E00305	C018	KL- team event Jan-29	18229	2094	16135
E00309	C008	IM- others Feb-2	8311	2305	6006
E00312	C014	CJ- concerts Feb-5	12314	3170	9144
E00314	C012	CI- birthday Feb-7	7373	2107	5266
E00315	C032	EG- others Feb-8	6787	2967	3820
E00319	C021	NL- conference Feb-12	13341	2466	10875
E00321	C001	QP- party Feb-14	11773	2702	9071
E00323	C021	IL- conference Feb-16	17764	3068	14696
E00324	C014	TG- conference Feb-17	16134	2492	13642
E00325	C022	AP- others Feb-18	8136	1491	6645
E00327	C013	GI- others Feb-20	14732	2143	12589
E00331	C005	BN- conference Feb-24	14520	3146	11374
E00332	C003	OV- others Feb-25	13880	2571	11309
E00334	C009	SI- birthday Feb-27	18246	4266	13980
E00338	C011	JB- others Mar-2	8427	3625	4802

Inference: Our customer does not have a view of his profit and this report will help him analyze which events were profitable for him along with the revenue and cost analysis. This gives him a wider view of his finances.

1.9.5 Customer Analysis:

Query 11

Business case: Finding the types of customers which return to our client

```
4   -- Customer Query 1 : Types and frequency of events by quarters --
5   select Type, Quarter_column, count(*) as Freq_of_events
6   from
7   (select event_table.* , case
8   when left(date,2) in (10) then "Q4"
9   when left(concat(0,date),2) in (01,02,03) then "Q1"
10  when left(concat(0,date),2) in (04,05,06) then "Q2"
11  when left(concat(0,date),2) in (07,08,09) then "Q3"
12  end as Quarter_column,
13  Customer.Org_CustName,Ind_CustFName,Ind_CustLName
14  from event event_table
15  left join
16  (select super_table.* , org.CompName as Org_CustName , ind.CustFName as Ind_CustFName,
17  ind.CustLName as Ind_CustLName, ind.CustAge as Ind_CustAge
18  from customer_super super_table
19  left join customerorg_sub org on super_table.CustID = org.CustID
20  left join customerind_sub ind on super_table.CustID = ind.CustID) Customer
21  on event_table.CustID = Customer.CustID) event_customer
22  group by Type, Quarter_column
23  order by Type, Freq_of_events desc;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Type	Quarter_column	Freq_of_events
birthday	Q3	10
birthday	Q2	4
birthday	Q1	3
birthday	Q4	1
concerts	Q3	6
concerts	Q2	5

Inference: The most frequent events are Conferences followed by Birthdays. This is consistent in both types of customers.

Query 12

Business Case: The types of clients and events which have more frequency

```
-- Customer Query 2 : Types and frequency of events by quarters breakdown by customers --
select Type, Quarter_column, count(*) as Freq_of_events,
case
when Org_CustName is not null then "Organization"
else "Individual" end as Customer_Type
from
(select event_table.* , case
when left(date,2) in (10) then "Q4"
when left(concat(0,date),2) in (01,02,03) then "Q1"
when left(concat(0,date),2) in (04,05,06) then "Q2"
when left(concat(0,date),2) in (07,08,09) then "Q3"
end as Quarter_column,
Customer.Org_CustName,Ind_CustFName,Ind_CustLName
from event event_table
left join
(select super_table.* , org.CompName as Org_CustName , ind.CustFName as Ind_CustFName,
ind.CustLName as Ind_CustLName, ind.CustAge as Ind_CustAge
from customer_super super_table
left join customerorg_sub org on super_table.CustID = org.CustID
left join customerind_sub ind on super_table.CustID = ind.CustID) Customer
on event_table.CustID = Customer.CustID) event_customer
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Type	Quarter_column	Freq_of_events	Customer_Type
birthday	Q3	7	Individual
birthday	Q1	3	Individual
birthday	Q2	3	Individual
birthday	Q3	3	Organization
birthday	Q2	1	Organization

Result 3

Inference: The Individual customers are the most frequent with 60% of the total events leaving Organizations with 40% of the total events.

1.9.6 Seasonality:

Query 13

Business Case: We need to see the busiest time of the year for our client

```
-- Seasonality Query 1 : Busiest time of the year --
select Quarter_column, Month_column, Type, count(*) as Freq_of_events
from
(select * , case
when left(date,2) in (10) then "October"
when left(concat(0,date),2) in (01) then "January"
when left(concat(0,date),2) in (02) then "February"
when left(concat(0,date),2) in (03) then "March"
when left(concat(0,date),2) in (04) then "April"
when left(concat(0,date),2) in (05) then "May"
when left(concat(0,date),2) in (06) then "June"
when left(concat(0,date),2) in (07) then "July"
when left(concat(0,date),2) in (08) then "August"
when left(concat(0,date),2) in (09) then "September"
end as Month_column,
case
when left(date,2) in (10) then "Q4"
when left(concat(0,date),2) in (01,02,03) then "Q1"
when left(concat(0,date),2) in (04,05,06) then "Q2"
when left(concat(0,date),2) in (07,08,09) then "Q3"
end as Quarter_column
from event ) universe
group by Quarter_column, Month_column, Type;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Quarter_column	Month_column	Type	Freq_of_events
Q1	February	birthday	2
Q1	February	concerts	1
Q1	February	conference	4
Q1	February	others	5
Q1	February	party	1
Q1	January	others	1

Result 4

Inference: The busiest time of the year is Quarter 3 with an even distribution of events in all the three months. Though, we see April having the maximum number of events in the year

Query 14

Business Case: We need to see the Total Revenue by Month and event type for our client

```sql
-- Seasonality Query 2 : Total Revenue by month and event type--
select Quarter_column, Month_column, Type, sum(charge_to_cust) as Total_Revenue
from
(select * , case
when left(date,2) in (10) then "October"
when left(concat(0,date),2) in (01) then "January"
when left(concat(0,date),2) in (02) then "February"
when left(concat(0,date),2) in (03) then "March"
when left(concat(0,date),2) in (04) then "April"
when left(concat(0,date),2) in (05) then "May"
when left(concat(0,date),2) in (06) then "June"
when left(concat(0,date),2) in (07) then "July"
when left(concat(0,date),2) in (08) then "August"
when left(concat(0,date),2) in (09) then "September"
end as Month_column,
case
when left(date,2) in (10) then "Q4"
when left(concat(0,date),2) in (01,02,03) then "Q1"
when left(concat(0,date),2) in (04,05,06) then "Q2"
when left(concat(0,date),2) in (07,08,09) then "Q3"
end as Quarter_column
from event ) universe
group by Quarter_column, Month_column, Type;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Quarter_column	Month_column	Type	Total_Revenue
Q1	February	birthday	25619
Q1	February	concerts	12314
Q1	February	conference	61759
Q1	February	others	51846
Q1	February	party	11773
Q1	January	others	17481

Result 5

Inference: Accordingly, the maximum revenue is generated in Q3 followed by Q2. Birthdays generate the maximum revenue for client.

1.10 Future Scope

The next crucial step is to take the Database to server such as MongoDB. It will be easier to store and retrieve when client has accumulated colossal amounts of data.

To make more data driven insights, we will collect more attributes to entities in the most granular form. This will help our client to understand his market better. As we are capturing the nascent stage of the company, we would like to implement few data quality checks to improve the quality of insights as well.

Since the client wants to review his business regularly, we will bring automation by connecting the SQL query outputs to a data visualization tool such as Tableau or Power BI, which will automatically refresh client's report in a timely manner.

This SQL project can be expanded to a Dashboard with Suggestion functionality where user can input the requirements and get optimal combinations of vendor/ staff suggestions to maximize profits.