




PMG SQL ASSESSMENT

Rohan Shukla
rohanshukla2998@gmail.com



Loading the Data in SQL

- SQL Query:
 - **Creating store_revenue table from CSV**

```
create table store_revenue (date date,  
brand_id int, store_location varchar(250),  
revenue float);
```
 - COPY store_revenue
FROM
'/Users/rohan/Downloads/store_revenue.csv'
DELIMITER ','
CSV HEADER;
 - **Creating marketing_data table from CSV**
 - create table marketing_data (date date, geo
varchar(2), impressions float, clicks float);
 - COPY marketing_data
FROM
'/Users/rohan/Downloads/marketing_data.csv'
DELIMITER ','
CSV HEADER;
- 



QUESTION 1

	sum_of_clicks	
	double precision	🔒
1	1792	

- SQL Query:

```
select sum(m.clicks) from  
marketing_data m;
```





QUESTION 2

- SQL Query:

```
select s.store_location, sum(s.revenue)
as revenue from store_revenue s group
by s.store_location order by 2 desc;
```

	store_location character varying (250)	revenue double precision
1	United States-CA	235237
2	United States-NY	51984
3	United States-TX	9629

QUESTION 3

	date date	subtring text	impressions double precision	clicks double precision	revenue double precision
1	2016-01-01	TX		2532	45
2	2016-01-01	CA		3425	63
3	2016-01-01	NY		3532	25
4	2016-01-01	MN		1342	784
5	2016-01-02	CA		1354	53
6	2016-01-02	NY		4643	85
7	2016-01-02	MN		2366	85
8	2016-01-02	TX		3643	23
9	2016-01-03	TX		2353	57
10	2016-01-03	MN		5783	87
11	2016-01-03	NY		4735	63
12	2016-01-03	CA		5258	36
13	2016-01-04	MN		9345	24
14	2016-01-04	NY		4754	36
15	2016-01-04	TX		5783	47
16	2016-01-04	CA		7854	85
17	2016-01-05	TX		2535	63
18	2016-01-05	MN		3452	25
19	2016-01-05	NY		2364	33
20	2016-01-05	CA		4678	73
21	2016-01-06	TX		[null]	[null]

- SQL Query:

```
select distinct s.date,  
SUBSTRING(s.store_location,  
15),AVG(m.impressions)::float as  
impressions, AVG(m.clicks)::float as  
clicks, SUM(s.revenue) as revenue from  
store_revenue s  
left outer join marketing_data m  
on m.date = s.date and m.geo =  
SUBSTRING(s.store_location, 15) group by  
s.store_location, s.date  
UNION  
select distinct m.date,  
m.geo,AVG(m.impressions)::float as  
impressions, AVG(m.clicks)::float as  
clicks, SUM(s.revenue) as revenue from  
store_revenue s  
right outer join marketing_data m  
on m.date = s.date and m.geo =  
SUBSTRING(s.store_location, 15) group by  
m.geo, m.date order by 1;
```

QUESTION 4

	geo character varying (2)	revenue_per_click_through double precision	click_through_rate double precision
1	CA	17126012.42903226	0.013735655102131242
2	NY	4302213.024793388	0.012083083682844019
3	TX	690040.8340425532	0.0139498990858364

Best Performing Store is California
Store with best Click Through Rate is Texas

Metric Used to Evaluate:

Click through rate proportional to Revenue
Conversion rate of clicks proportional to Revenue
Assuming product costs common for all stores

Thus,

$CTR * \text{Conv. Rate} * \text{Cost} \sim \text{Revenue}$

$\text{Conv. Rate} \sim \text{Revenue} / \text{CTR}$

Hence, CA has the best Conversion Rate of Clicks,
making it the most efficient store

- SQL Query:


```
with click_revenue as (  
  select distinct m.date,  
    m.geo, AVG(m.impressions)::float as  
    impressions, AVG(m.clicks)::float as  
    clicks, SUM(s.revenue) as revenue from  
    store_revenue s  
  join marketing_data m  
    on m.date = s.date and m.geo =  
    SUBSTRING(s.store_location, 15) group  
    by m.geo, m.date)  
select c.geo,  
  SUM(c.revenue) / (SUM(c.clicks) / SUM(c.imp  
    ressions)) as  
  revenue_per_click_through,  
  SUM(c.clicks) / SUM(c.impressions) as  
  click_through_rate from click_revenue c  
group by c.geo order by 2 desc;
```



QUESTION 5

	states text	revenue double precision
1	CA	235237
2	NY	51984
3	TX	9629

- SQL Query:

```
select SUBSTRING(s.store_location,15)
as states, SUM(s.revenue)::float as
revenue from store_revenue s
group by s.store_location order by 2
desc limit 10;
```
- 



THANK YOU