

Tình hình Thị trường và Hiệu suất Quảng cáo trên Website ABC



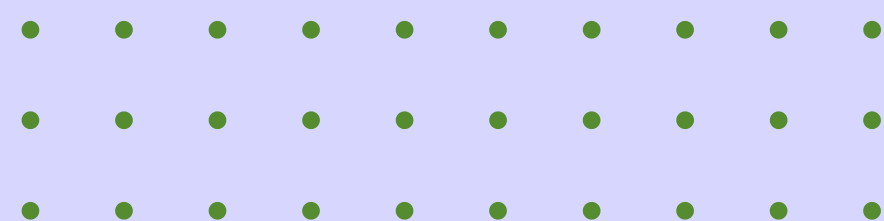
01. Giới thiệu Dataset

02. Công cụ sử dụng

03. Các bước thực hiện

04. Phân tích

05. Đề xuất giải pháp



Mục Lục



Giới thiệu dataset

Công ty X có sản phẩm chủ đạo là **website ABC**, là website tìm kiếm và so sánh booking khách sạn lớn thuộc top đầu thế giới. Team Marketplace là một team lớn gồm các data analysts, data scientists đến từ nhiều quốc gia khác nhau. Nhiệm vụ của team là cung cấp các phân tích dữ liệu nhằm hỗ trợ việc đưa ra quyết định của công ty X.

Dữ liệu được cho là daily performance của **3 nhà quảng cáo (advertisers) A, B và C** trên website ABC trong năm 2019 tại 1 quốc gia



Giới thiệu dataset

date	ttt_group	clicks_A	clicks_B	clicks_C	cost_A	cost_B	cost_C	bookings_A	bookings_B	bookings_C	booking_rev_A	booking_rev_B	booking_rev_C
2019-01-01	short	110658	62984	23897	107515	57033	19170	5359	2440	1157	815380	374662	173965
2019-01-01	medium	61201	8115	81676	47501	4893	65722	2373	252	3167	371986	39742	512900
2019-01-01	long	4945	7970	301992	2378	4068	242810	118	153	7231	19093	23636	1190356
2019-01-02	short	99122	54777	21459	96726	49634	17295	4798	2121	1039	720475	328767	164653
2019-01-02	medium	59130	8040	79556	45954	4869	64165	2294	250	3086	360364	39281	474731
2019-01-02	long	4202	6524	284020	2022	3317	231466	101	125	6807	15838	18765	1138882
2019-01-03	short	86154	49526	19704	83567	44827	15894	4168	1917	953	642517	294179	158464
2019-01-03	medium	57538	7733	76565	44838	4689	61834	2234	240	2972	339218	38368	445813
2019-01-03	long	4409	6536	263483	2121	3303	211347	106	125	6319	16421	19368	1021024
2019-01-04	short	75563	43967	17867	73186	39797	14451	3654	1701	864	558728	260389	140913
2019-01-04	medium	55124	7097	74045	43154	4300	60145	2142	221	2877	348974	33317	446521
2019-01-04	long	3789	5502	244675	1822	2772	198101	91	106	5872	14031	16632	886951
2019-01-05	short	64958	39925	16296	62977	36422	13291	3139	1543	788	508980	239212	125799
2019-01-05	medium	55259	6576	68129	43446	3962	54985	2149	205	2649	342436	31326	409411
2019-01-05	long	3662	4992	224717	1768	2506	181524	88	96	5395	14526	15117	848900
2019-01-06	short	117272	68608	27152	113299	61987	21848	5664	2651	1311	864767	423745	197976
2019-01-06	medium	96050	12415	137291	74761	7483	111740	3738	386	5343	565066	60799	812847
2019-01-06	long	7060	9568	408676	3407	4796	327505	170	184	9813	27480	27867	1506297
2019-01-07	short	97953	55346	22733	94942	49950	18357	4728	2137	1097	758496	321490	174453
2019-01-07	medium	91774	11949	113427	71704	7236	91000	3575	372	4418	576164	56522	751332
2019-01-07	long	5813	8461	348046	2805	4278	280077	140	163	8358	22476	26444	1352270

- Thời gian
 - Phân khúc khách hàng
 - Số lượng clicks người dùng thực hiện
 - Số tiền mỗi nhà quảng cáo phải trả cho website ABC
- Số lượng booking khách sạn được thực hiện bởi người dùng thông qua website ABC
 - Số tiền mà người dùng tiêu cho booking khách sạn (tức chính là gross revenue của mỗi nhà quảng cáo)

Công cụ sử dụng



Python



PowerBI

Các bước thực hiện

```
1 from google.colab import files
2 import pandas as pd
3 upload = files.upload()

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.
Saving marketplace_data_2019(1).xlsx to marketplace_data_2019(1).xlsx

[ ] 1 df_marketplace19 = pd.read_excel('marketplace_data_2019(1).xlsx')
    2 df_marketplace19.head(5)
```

	date	ttt_group	clicks_A	clicks_B	clicks_C	cost_A	cost_B	cost_C	bookings_A	bookings_B	bookings_C	booking_rev_A	booking_rev_B
0	2019-01-01	long	4945	7970	301992	2378	4068	242810	118	153	7231	19093	23636
1	2019-01-01	short	110658	62984	23897	107515	57033	19170	5359	2440	1157	815380	374662
2	2019-01-01	medium	61201	8115	81676	47501	4893	65722	2373	252	3167	371986	39742
3	2019-01-02	long	4202	6524	284020	2022	3317	231466	101	125	6807	15838	18765
4	2019-01-02	short	99122	54777	21459	96726	49634	17295	4798	2121	1039	720475	328767

- Sử dụng thư viện pandas để đọc file trên Python.
- Check 5 dòng đầu tiên của dữ liệu

```
1 df_marketplace19.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1095 entries, 0 to 1094
Data columns (total 14 columns):
#   Column          Non-Null Count  Dtype  
---  -
0   date             1095 non-null  datetime64[ns]
1   ttt_group        1095 non-null  object  
2   clicks_A         1095 non-null  int64   
3   clicks_B         1095 non-null  int64   
4   clicks_C         1095 non-null  int64   
5   cost_A           1095 non-null  int64   
6   cost_B           1095 non-null  int64   
7   cost_C           1095 non-null  int64   
8   bookings_A       1095 non-null  int64   
9   bookings_B       1095 non-null  int64   
10  bookings_C       1095 non-null  int64   
11  booking_rev_A    1095 non-null  int64   
12  booking_rev_B    1095 non-null  int64   
13  booking_rev_C    1095 non-null  int64   
dtypes: datetime64[ns](1), int64(12), object(1)
memory usage: 119.9+ KB
```

Sử dụng lệnh .info để xem thông tin và kiểm tra các dữ liệu null

Các bước thực hiện

fx

= Table.TransformColumnTypes("#Promoted Headers",{{"date", type date}, {"ttt_group", type text}, {"clicks_A", Int64.Type},

	date	ttt_group	clicks_A	clicks_B	clicks_C	cost_A
	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>
1	1/6/2019	short	117272	68608	27152	
2	1/6/2019	medium	96050	12415	137291	
3	1/6/2019	long	7060	9568	408676	
4	1/13/2019	short	109700	64097	25477	
5	1/13/2019	medium	128847	16787	172639	
6	1/13/2019	long	6473	9551	415381	
7	1/20/2019	short	132562	83019	31839	
8	1/20/2019	medium	179077	22359	196475	
9	1/20/2019	long	5544	9485	422090	
10	1/27/2019	short	189934	118627	44164	
11	1/27/2019	medium	214657	25781	238325	
12	1/27/2019	long	5555	9852	428138	
13	2/3/2019	short	275571	170168	64992	
14	2/3/2019	medium	265241	26041	262481	
15	2/3/2019	long	4873	10485	436051	
16	2/10/2019	short	387588	231434	97422	
17	2/10/2019	medium	294408	27456	292986	
18	2/10/2019	long	3774	9833	448219	

Query Settings

PROPERTIES

Name
marketplace_data_2019

All Properties

APPLIED STEPS

Source

Promoted Headers

Changed Type

Unpivoted Columns

Split Column by Delimiter

Changed Type1

Renamed Columns

Reordered Columns

Pivoted Column

- Load dữ liệu vào PowerBI và bắt đầu các bước Transform dữ liệu

fx

= Table.UnpivotOtherColumns("#Changed Type", {"date", "ttt_group"}, "Attribute", "Value")

	date	ttt_group	Attribute	Value
	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>
1	1/6/2019	short	clicks_A	117272
2	1/6/2019	short	clicks_B	68608
3	1/6/2019	short	clicks_C	27152
4	1/6/2019	short	cost_A	113299
5	1/6/2019	short	cost_B	61987
6	1/6/2019	short	cost_C	21848
7	1/6/2019	short	bookings_A	5664
8	1/6/2019	short	bookings_B	2651
9	1/6/2019	short	bookings_C	1311
10	1/6/2019	short	booking_rev_A	864767
11	1/6/2019	short	booking_rev_B	423745
12	1/6/2019	short	booking_rev_C	197976
13	1/6/2019	medium	clicks_A	96050
14	1/6/2019	medium	clicks_B	12415
15	1/6/2019	medium	clicks_C	137291
16	1/6/2019	medium	cost_A	74761
17	1/6/2019	medium	cost_B	7483
18	1/6/2019	medium	cost_C	111740

Query Settings

PROPERTIES

Name
marketplace_data_2019

All Properties

APPLIED STEPS

Source

Promoted Headers

Changed Type

Unpivoted Columns

Split Column by Delimiter

Changed Type1

Renamed Columns

Reordered Columns

Pivoted Column

- Giữ lại trường Date, TTT_group và Unpivot các trường còn lại

Các bước thực hiện

Query Editor

fx = Table.SplitColumn("#Unpivoted Columns", "Attribute", Splitter.SplitTextByEachDelimiter({"_"}, QuoteStyle.Csv, true),

	date	ttt_group	Attribute.1	Attribute.2	Value
Valid 100%	Valid 100%	Valid 100%	Valid 100%	Valid 100%	Valid 100%
Error 0%	Error 0%	Error 0%	Error 0%	Error 0%	Error 0%
Empty 0%	Empty 0%	Empty 0%	Empty 0%	Empty 0%	Empty 0%
1	1/6/2019	short	clicks	A	117272
2	1/6/2019	short	clicks	B	68608
3	1/6/2019	short	clicks	C	27152
4	1/6/2019	short	cost	A	113299
5	1/6/2019	short	cost	B	61987
6	1/6/2019	short	cost	C	21848
7	1/6/2019	short	bookings	A	5664
8	1/6/2019	short	bookings	B	2651
9	1/6/2019	short	bookings	C	1311
10	1/6/2019	short	booking_rev	A	864767
11	1/6/2019	short	booking_rev	B	423745
12	1/6/2019	short	booking_rev	C	197976
13	1/6/2019	medium	clicks	A	96050
14	1/6/2019	medium	clicks	B	12415
15	1/6/2019	medium	clicks	C	137291
16	1/6/2019	medium	cost	A	74761
17	1/6/2019	medium	cost	B	7483
18	1/6/2019	medium	cost	C	111740

Query Settings

PROPERTIES

Name
marketplace_data_2019

All Properties

APPLIED STEPS

- Source
- Promoted Headers
- Changed Type
- Unpivoted Columns
- Split Column by Delimiter**
- Changed Type1
- Renamed Columns
- Reordered Columns
- Pivoted Column

- Tách dữ liệu để lọc ra một cột mới chỉ gồm 3 nhà quảng cáo A, B và C

	date	ttt_group	Advertisers	clicks	cost	bookings
Valid	100%	Valid	100%	Valid	100%	Valid
Error	0%	Error	0%	Error	0%	Error
Empty	0%	Empty	0%	Empty	0%	Empty
1	1/1/2019	long	A	4945	2378	
2	1/1/2019	long	B	7970	4068	
3	1/1/2019	long	C	301992	242810	
4	1/1/2019	medium	A	61201	47501	
5	1/1/2019	medium	B	8115	4893	
6	1/1/2019	medium	C	81676	65722	
7	1/1/2019	short	A	110658	107515	
8	1/1/2019	short	B	62984	57033	
9	1/1/2019	short	C	23897	19170	
10	1/2/2019	long	A	4202	2022	
11	1/2/2019	long	B	6524	3317	
12	1/2/2019	long	C	284020	231466	
13	1/2/2019	medium	A	59130	45954	
14	1/2/2019	medium	B	8040	4869	
15	1/2/2019	medium	C	79556	64165	
16	1/2/2019	short	A	99122	96726	
17	1/2/2019	short	B	54777	49634	
18	1/2/2019	short	C	21459	17295	

- Pivot dữ liệu theo các advertisers A, B, C

Các bước thực hiện

✕ ✓		1 Date = CALENDAR(MIN(marketplace19[date]), MAX(marketplace19[date].[Date]))		Data >>	
Date ▾				Search	
1/1/2019				> Date	
1/2/2019				> marketplace19	
1/3/2019					
1/4/2019					
1/5/2019					
1/6/2019					
1/7/2019					
1/8/2019					
1/9/2019					
1/10/2019					
1/11/2019					
1/12/2019					
1/13/2019					
1/14/2019					
1/15/2019					
1/16/2019					
1/17/2019					
1/18/2019					
1/19/2019					
1/20/2019					
1/21/2019					

- Tạo thêm bảng Date từ hàm Calendar và Min, Max

Các bước thực hiện

([DAX])

```
1 BookingConversion = DIVIDE([TotalBookings],[TotalClicks],0)
```

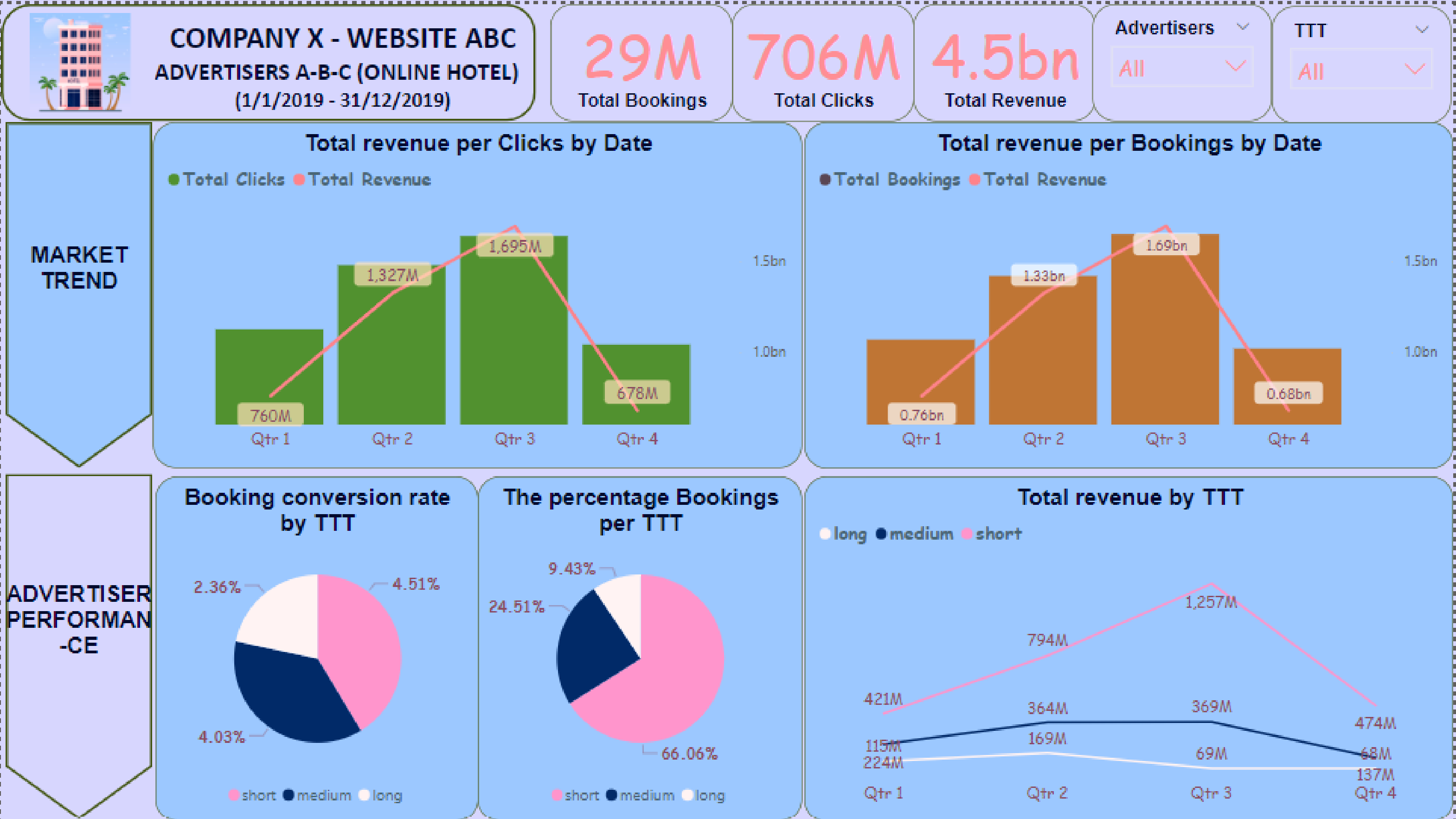
```
1 CostPerBookings = DIVIDE([TotalCost],[TotalBookings],0)
```

```
1 CostPerClicks = DIVIDE([TotalCost],[TotalClicks],0)
```

```
1 ROI =  
2 DIVIDE(  
3     SUMX(marketplace19,'marketplace19'[Booking_rev] - 'marketplace19'[Cost]),  
4     SUMX(marketplace19,'marketplace19'[Cost]),  
5     0  
6 )
```

```
1 TotalProfit = SUMX('marketplace19', 'marketplace19'[Booking_rev]) * 0.15
```

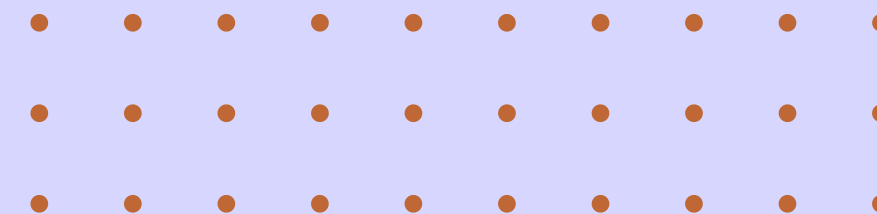
Phân tích vấn đề: Market Trend



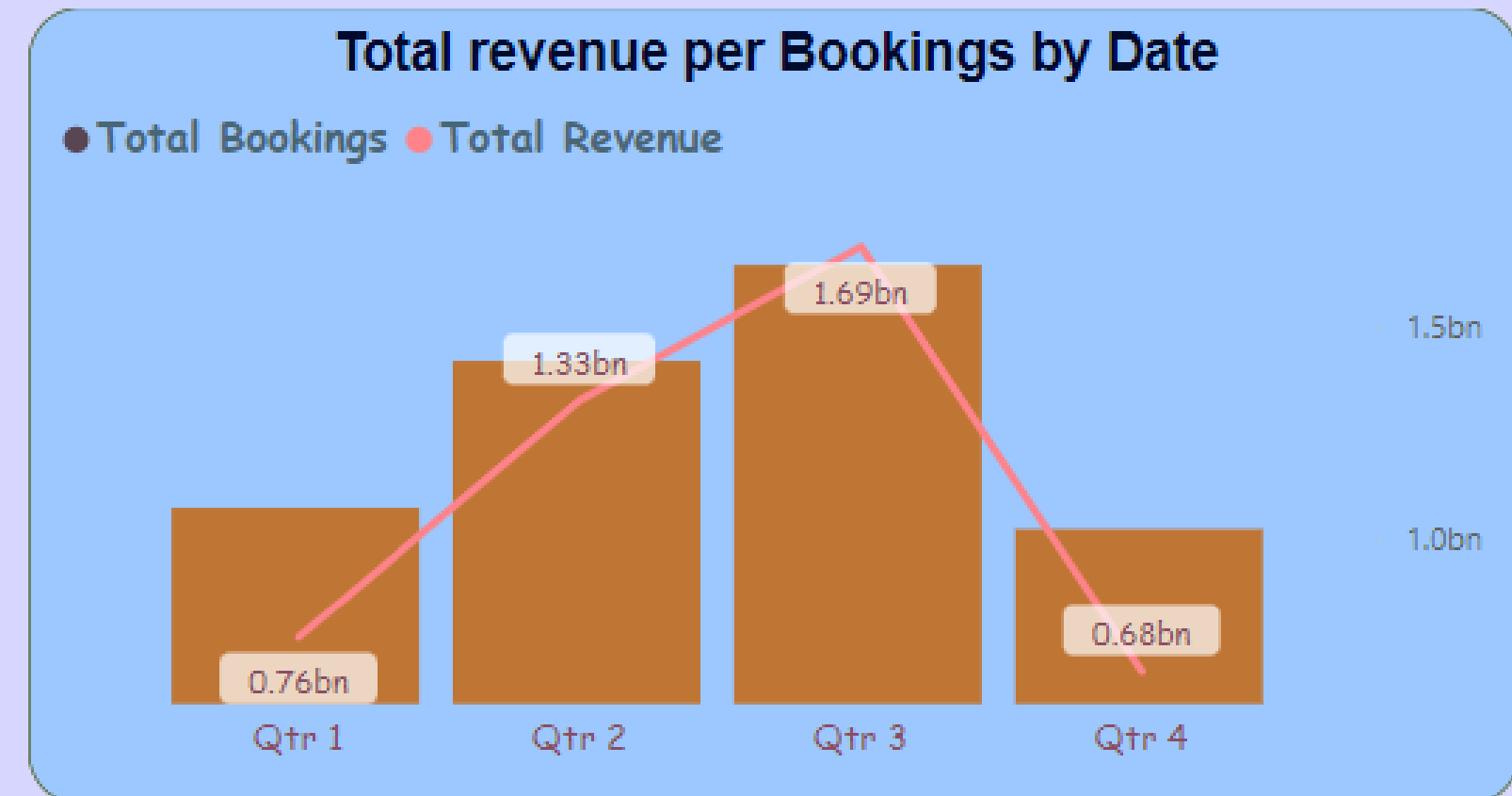
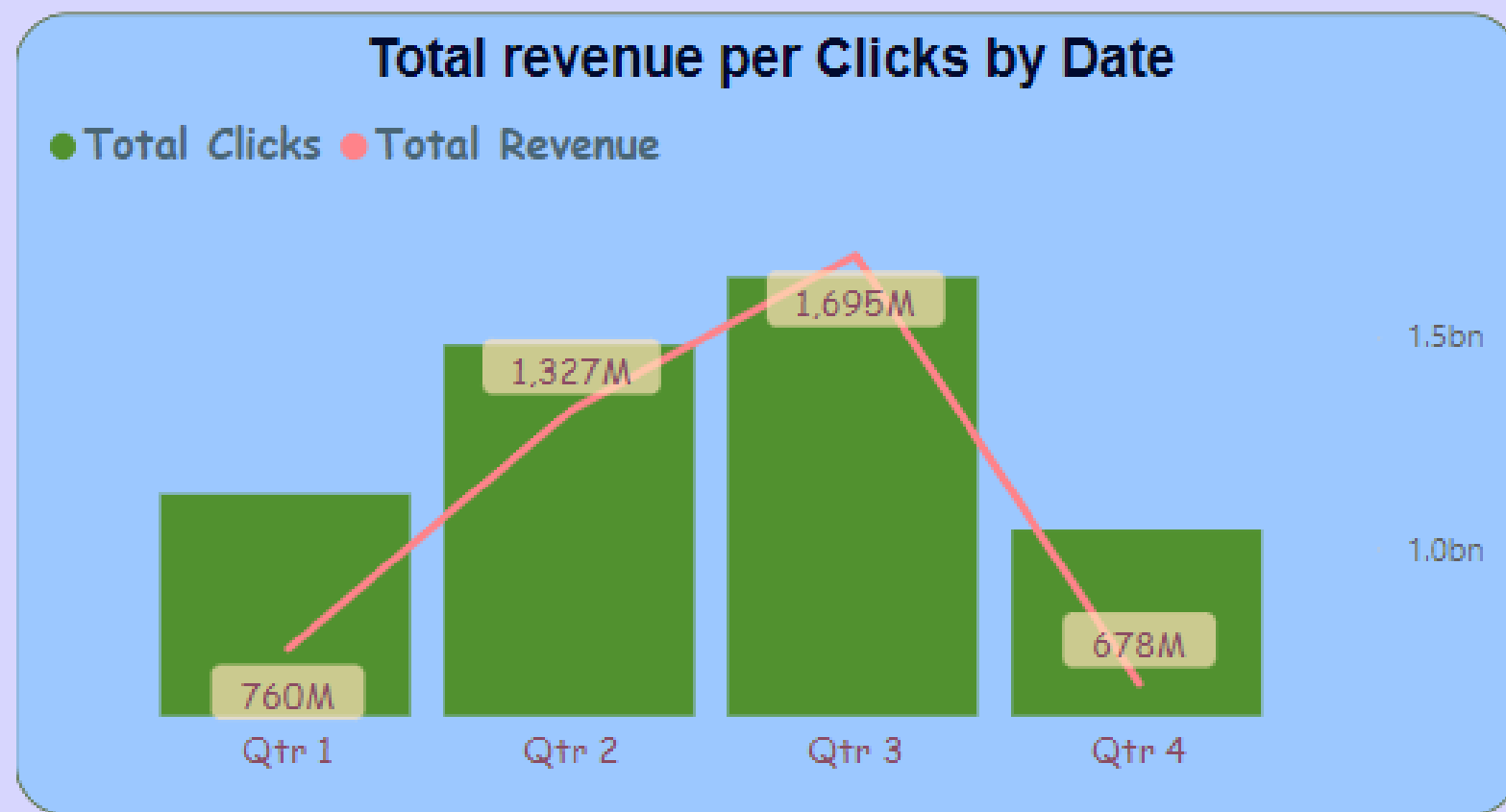


MARKET TREND

Hành vi sử dụng dịch vụ đặt phòng của khách hàng
trên website ABC

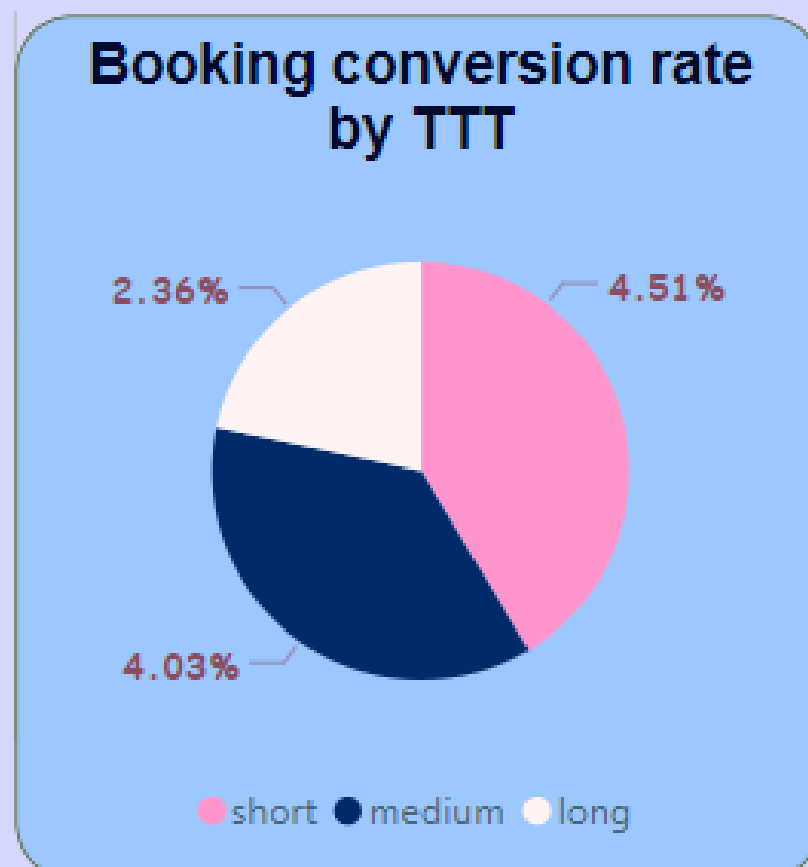


Phân tích vấn đề: Market Trend



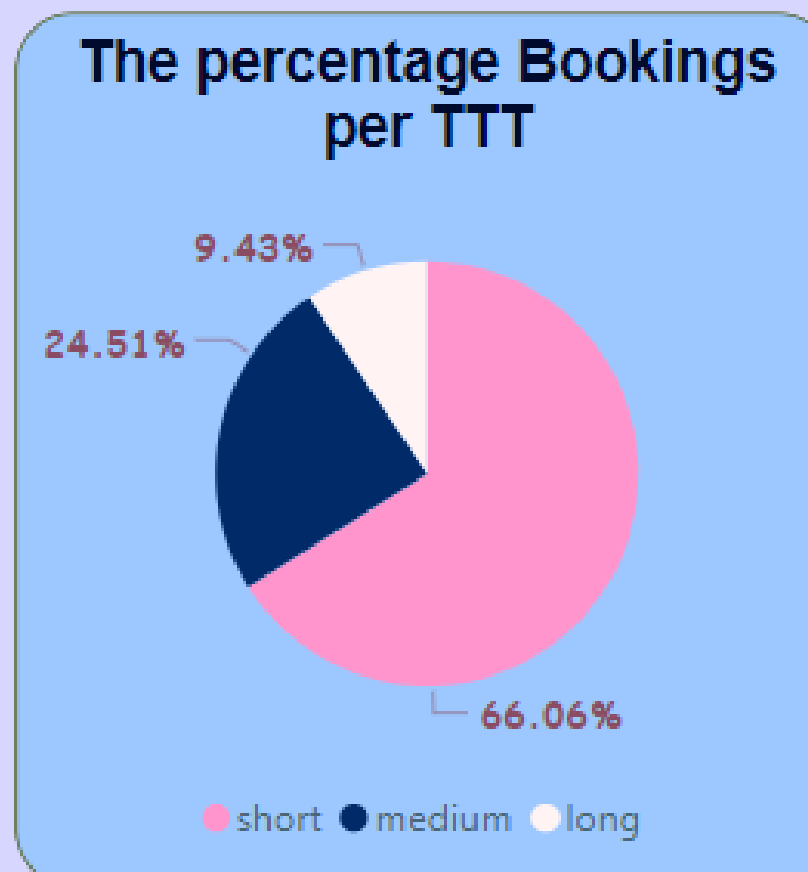
- Từ biểu đồ 1 và biểu đồ 2 có thể thấy total revenue, total clicks và total bookings có **mối tương quan đồng biến** với nhau. Cả 3 đều tăng dần đều từ quý 1 năm 2019 và **đạt đỉnh ở quý 3** cùng năm rồi giảm dần khi đến quý 4.
- Từ thông tin trên có thể đoán đây là dữ liệu của **Việt Nam**. Bởi vì 3 quý đầu năm Việt Nam có nhiều ngày nghỉ nên nhu cầu du lịch **tăng cao** và đặc biệt là vào quý 3 (tháng 6-tháng 9) đây là thời gian nghỉ hè của học sinh.

Phân tích vấn đề: Market Trend



Tỷ lệ booking conversion rate **cao** trong phân khúc **short và medium**, trong khi phân khúc **long** có tỷ lệ booking conversion rate **thấp hơn**

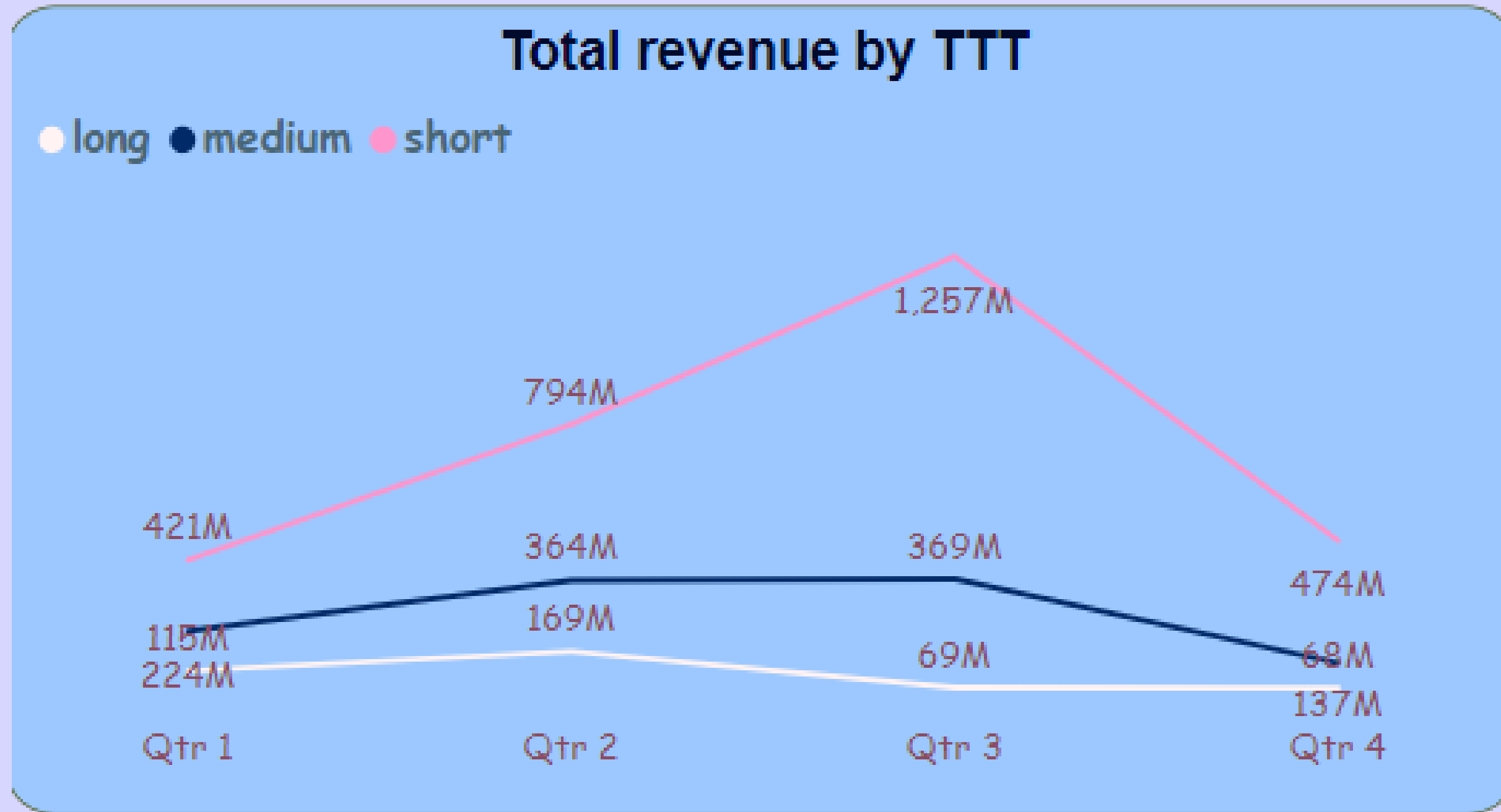
=> Những khách hàng có TTT ngắn và trung bình có xu hướng tìm được KH phù hợp và thực hiện việc đặt phòng khách sạn từ website ABC hơn khách hàng có TTT dài



Tỷ lệ phần trăm của số lượng booking trong mỗi phân khúc TTT. Khách hàng có **TTT ngắn** có xu hướng đặt phòng khách sạn **nhiều hơn** khách hàng có TTT medium và short (chiếm 2/3 trên tổng số). KH có TTT dài thì thường ít đặt phòng khách sạn trên website ABC hơn

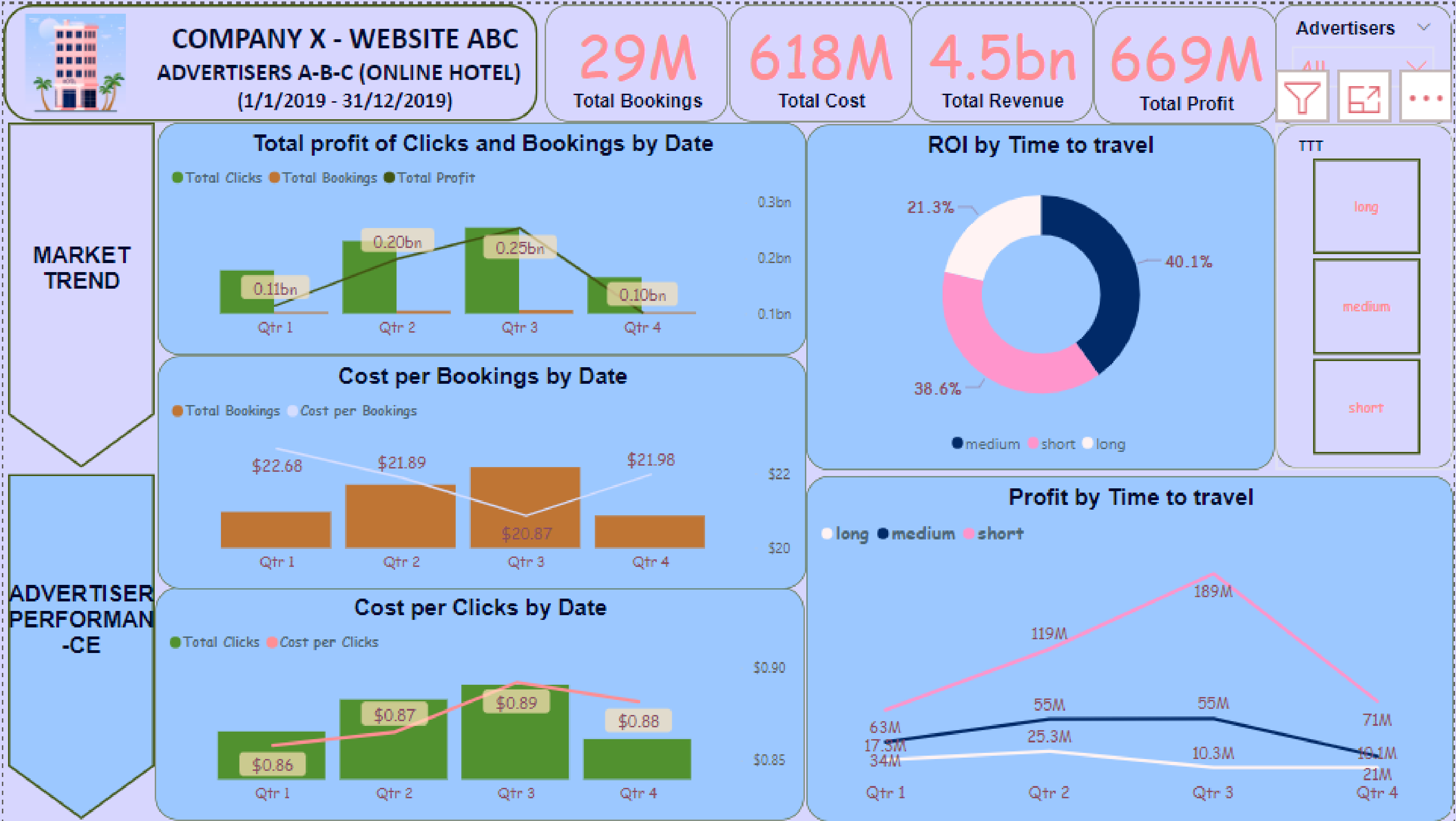
TTT (Time-to-Travel): số ngày kể từ ngày khách hàng đặt phòng khách sạn cho đến ngày khách hàng check-in tại khách sạn đã đặt

Phân tích vấn đề: Market Trend

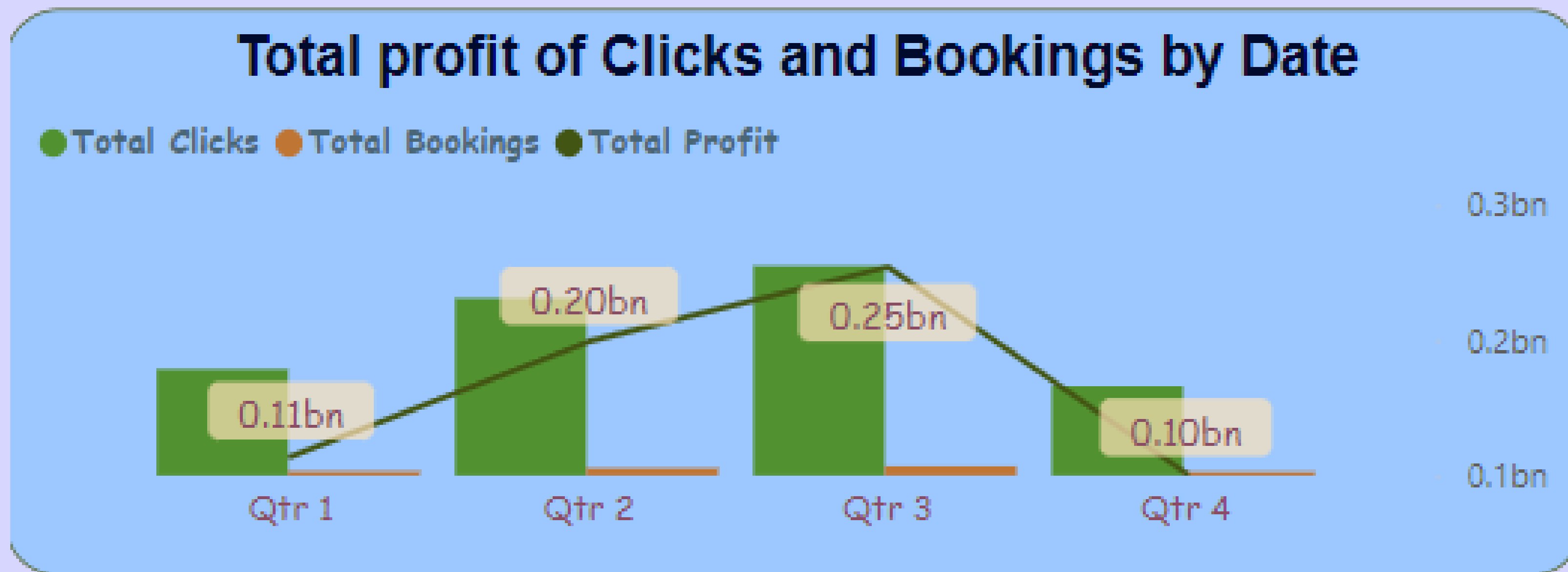


Cả phân khúc TTT short và medium đều cho doanh thu **cao nhất vào quý 3** và giảm dần khi về quý 4. Còn đối với TTT long thì doanh thu có sự tăng giảm nhẹ trong suốt cả năm và đạt doanh thu cao vào quý 2

Phân tích vấn đề: Advertiser Performance

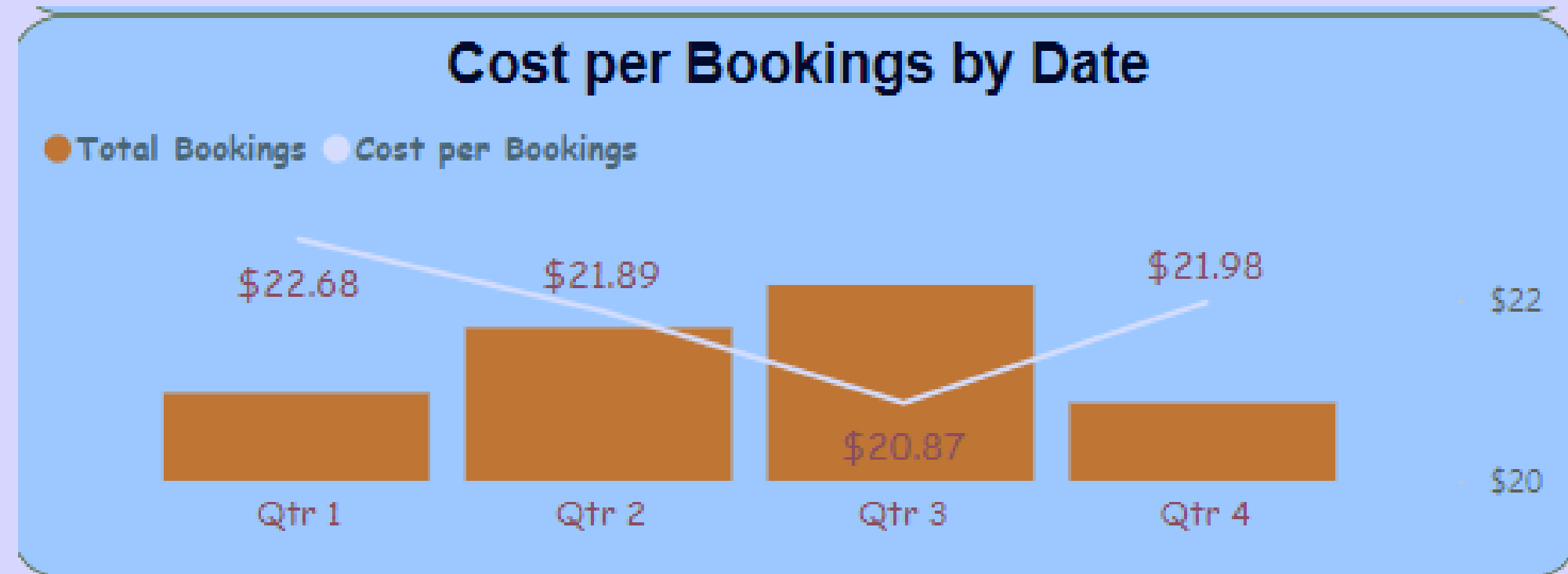


Phân tích vấn đề: Advertiser Performance

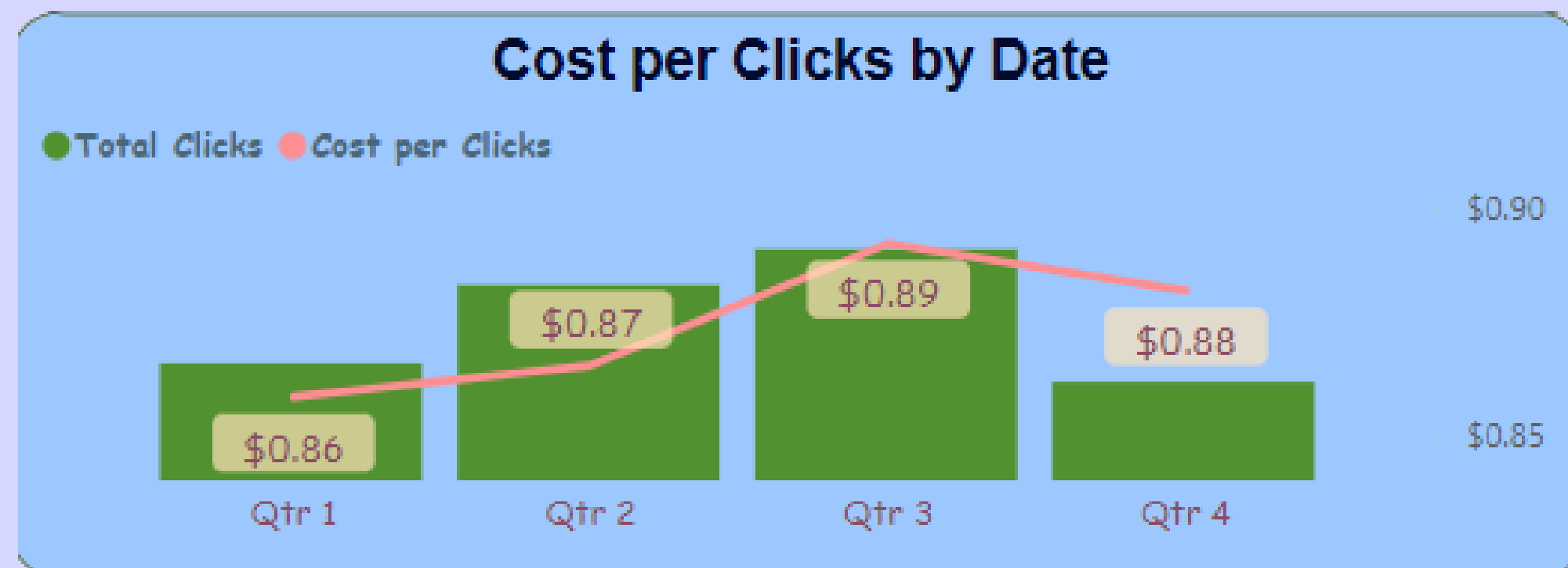


- Total Profit **tăng đều và mạnh** ở 3 quý đầu nhưng **giảm khá mạnh** ở quý 4.
- Mặc dù, số lượng clicks ở quý 4 vẫn ở mức tương đối nhưng lợi nhuận thu về **tỉ lệ nghịch** so với số lượng clicks và số lượng booking hầu như khá ít

Phân tích vấn đề: Advertiser Performance

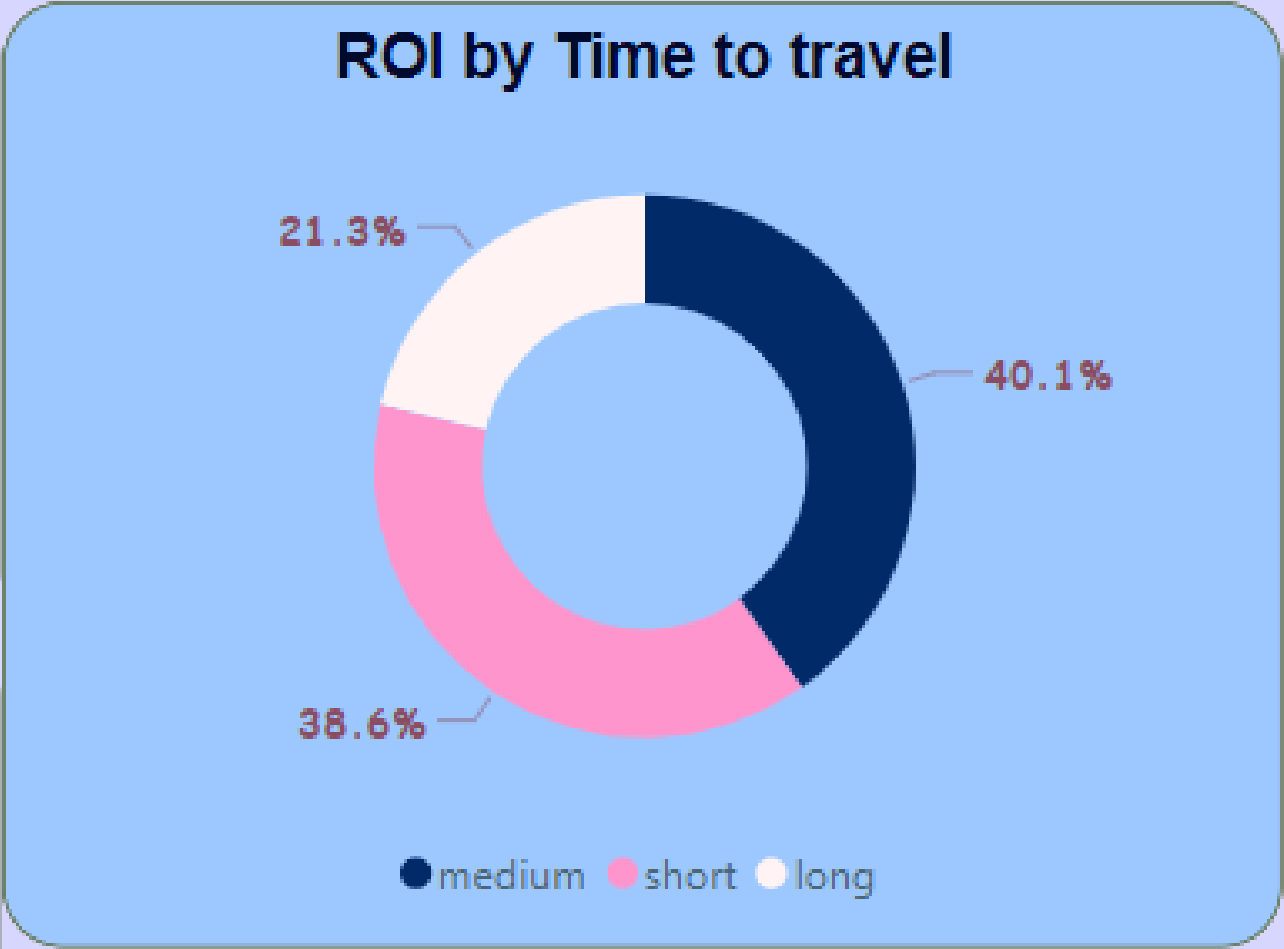


Tổng số lượng booking tăng dần từ quý 1 đến quý 3 trong khi cost per booking giảm trong cùng kỳ -> điều này cho thấy chiến dịch quảng cáo đang có hiệu quả tốt hơn và đạt được nhiều kết quả với chi phí ít hơn. Tuy nhiên đến quý 4 số lượng booking giảm nhưng cost per Booking lại tăng



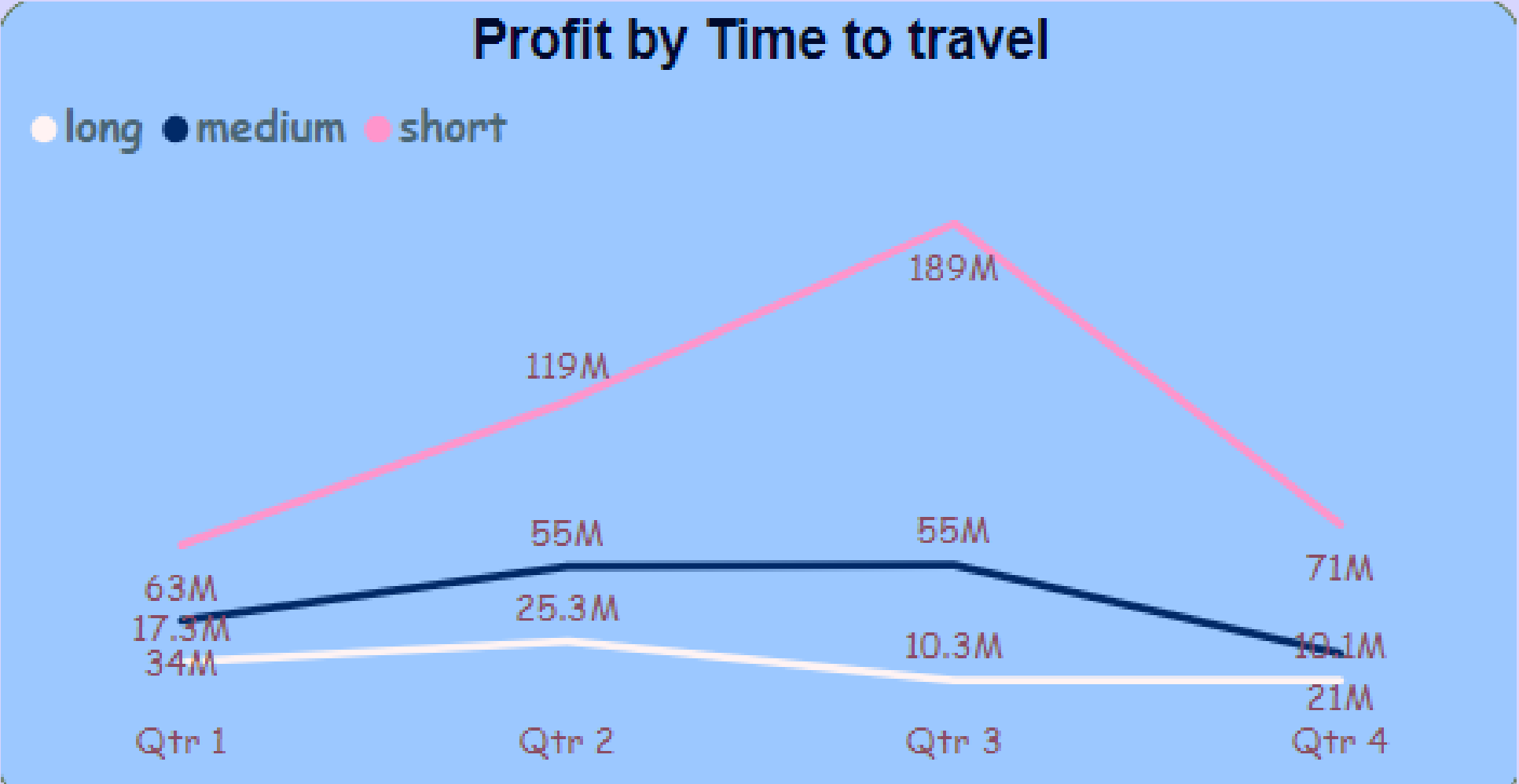
Tổng số lượng clicks tăng dần tỉ lệ thuận với cost per clicks, đến quý 4 cost per click giảm nhẹ nhưng totals clicks lại giảm đáng kể

Phân tích vấn đề: Advertiser Performance



→ Lợi nhuận và hiệu quả quảng cáo của mỗi nhà quảng cáo hầu như rơi vào nhóm SHORT

Lợi nhuận phần lớn đến từ TTT Short và đạt đỉnh ở quý 3 ←



Đề xuất giải pháp

Action 1

Tối ưu hóa chiến dịch quảng cáo: Cải thiện website như thiết kế giao diện thân thiện; giảm giá, đưa ra mức giá cạnh tranh hoặc các gói khuyến mãi

Action 2

Tập trung quảng cáo vào phân khúc **Short**

Action 3

Điều chỉnh chiến lược quảng cáo cho phân khúc **Medium**

Action 4

Đánh giá lại quảng cáo cho phân khúc **Long**: điều chỉnh đối tượng và nội dung quảng cáo để tăng khả năng chuyển đổi và sinh lợi từ phân khúc này

THANK YOU