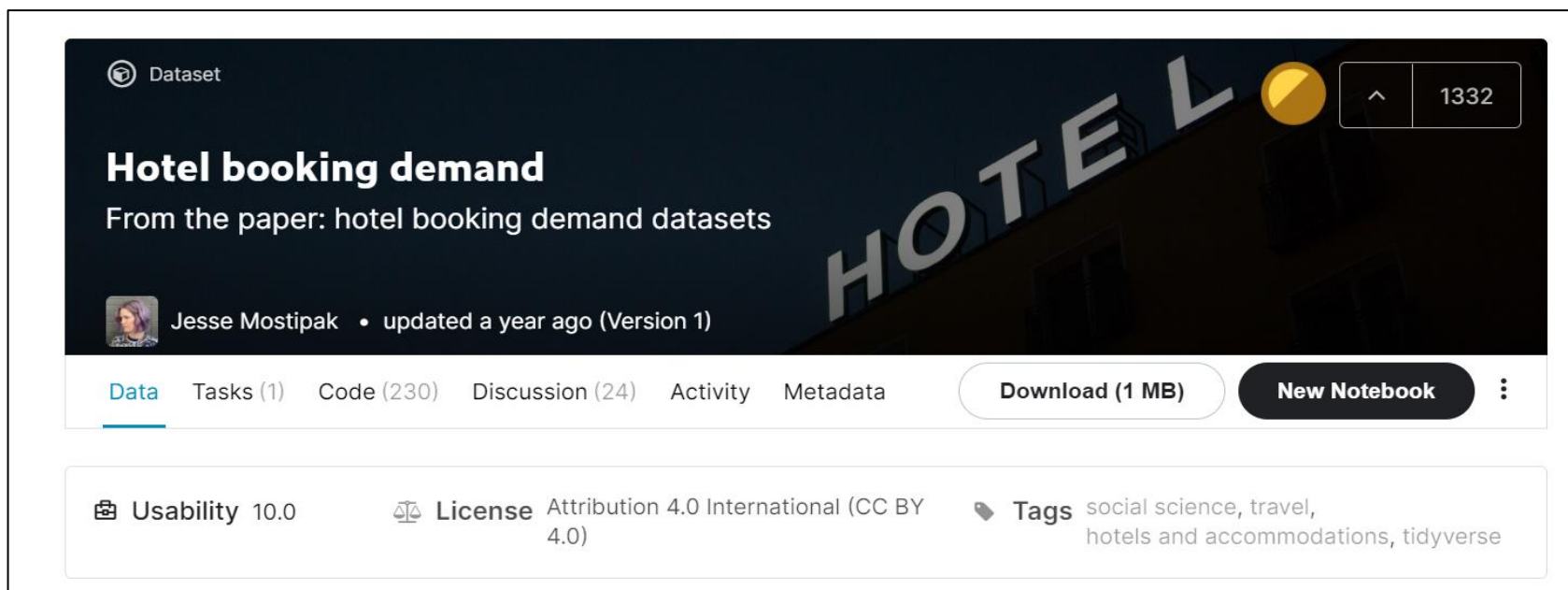


Capstone Project 3

Introduction

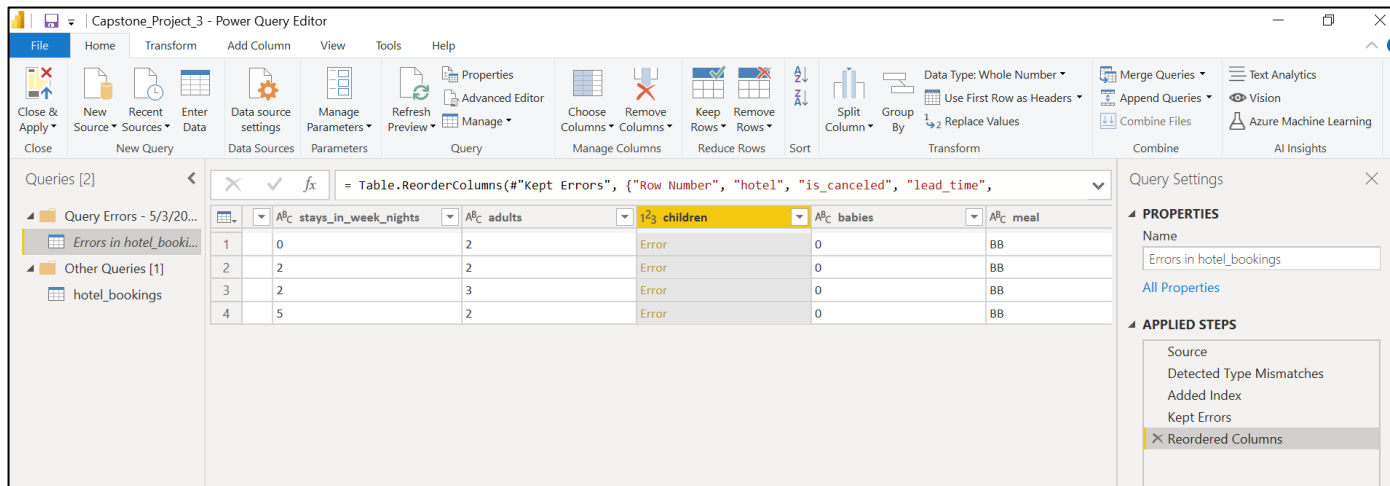
► Source : 



The screenshot shows the Kaggle dataset page for "Hotel booking demand". At the top, it says "Dataset" with a circular icon. The title "Hotel booking demand" is prominently displayed, followed by the subtitle "From the paper: hotel booking demand datasets". Below this, the creator's name "Jesse Mostipak" is shown with a small profile picture and a note that the dataset was "updated a year ago (Version 1)". The background of the header features a dark image of a building with "HOTEL" written on its facade. To the right of the title, there is a gold medal icon and a box showing an upward arrow and the number "1332". Below the header, there is a navigation bar with links for "Data", "Tasks (1)", "Code (230)", "Discussion (24)", "Activity", and "Metadata". To the right of these links are two buttons: "Download (1 MB)" and "New Notebook". At the bottom of the page, there is a section with "Usability 10.0", "License Attribution 4.0 International (CC BY 4.0)", and "Tags social science, travel, hotels and accommodations, tidyverse".

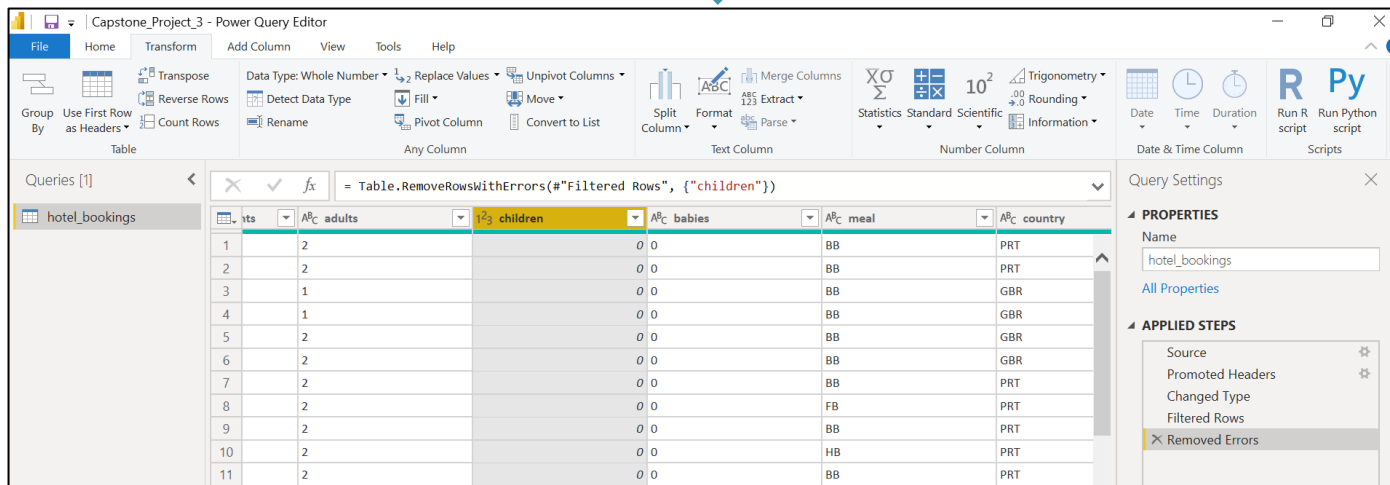
Data Preparations:

► Removed Errors



The screenshot shows the Power Query Editor interface for a query named "Errors in hotel_bookings". The formula bar displays the formula: `= Table.ReorderColumns(#"Kept Errors", {"Row Number", "hotel", "is_canceled", "lead_time",`. The data table has columns: `stays_in_week_nights`, `adults`, `children`, `babies`, and `meal`. The `children` column contains "Error" values for rows 1, 2, 3, and 4. The right-hand pane shows the "APPLIED STEPS" list, which includes "Reordered Columns".

	stays_in_week_nights	adults	children	babies	meal
1	0	2	Error	0	BB
2	2	2	Error	0	BB
3	2	3	Error	0	BB
4	5	2	Error	0	BB

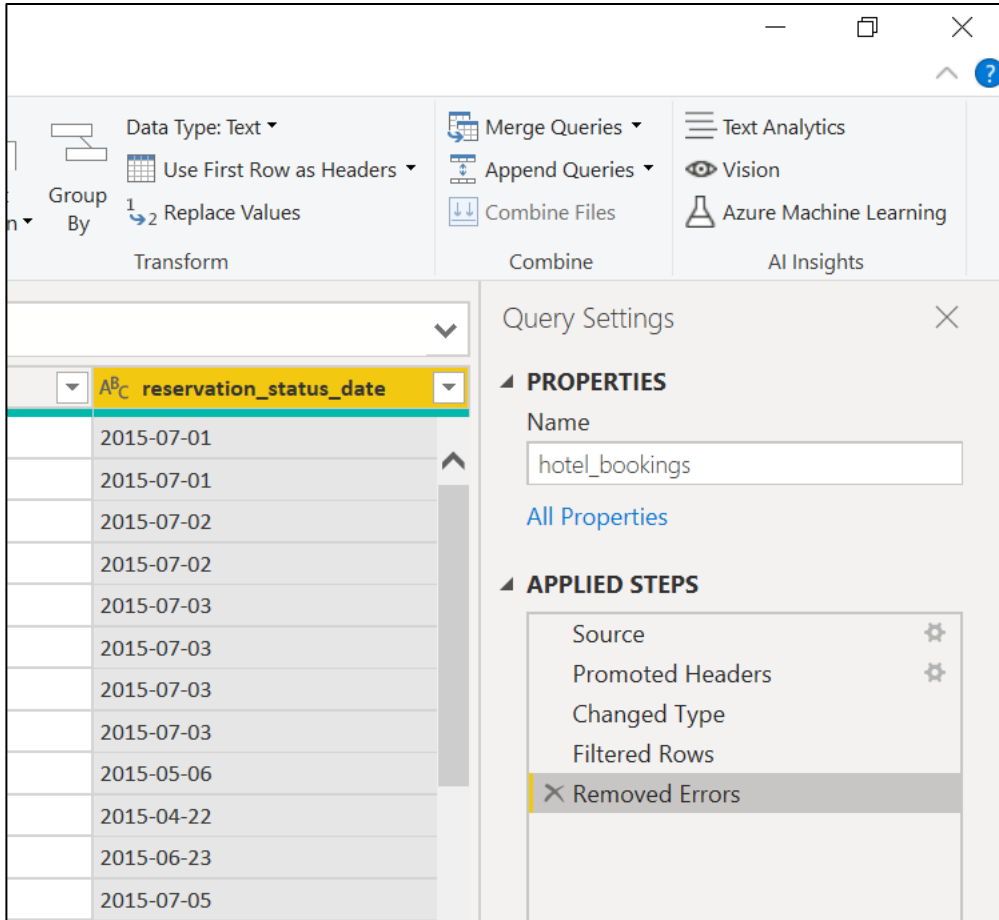


The screenshot shows the Power Query Editor interface for a query named "hotel_bookings". The formula bar displays the formula: `= Table.RemoveRowsWithErrors(#"Filtered Rows", {"children"})`. The data table has columns: `stays_in_week_nights`, `adults`, `children`, `babies`, `meal`, and `country`. The `children` column now contains numerical values (0 or 1) for all rows. The right-hand pane shows the "APPLIED STEPS" list, which includes "Removed Errors".

	stays_in_week_nights	adults	children	babies	meal	country
1	2	2	0	0	BB	PRT
2	2	2	0	0	BB	PRT
3	1	2	0	0	BB	GBR
4	1	2	0	0	BB	GBR
5	2	2	0	0	BB	GBR
6	2	2	0	0	BB	GBR
7	2	2	0	0	BB	PRT
8	2	2	0	0	BB	PRT
9	2	2	0	0	BB	PRT
10	2	2	0	0	BB	PRT
11	2	2	0	0	BB	PRT

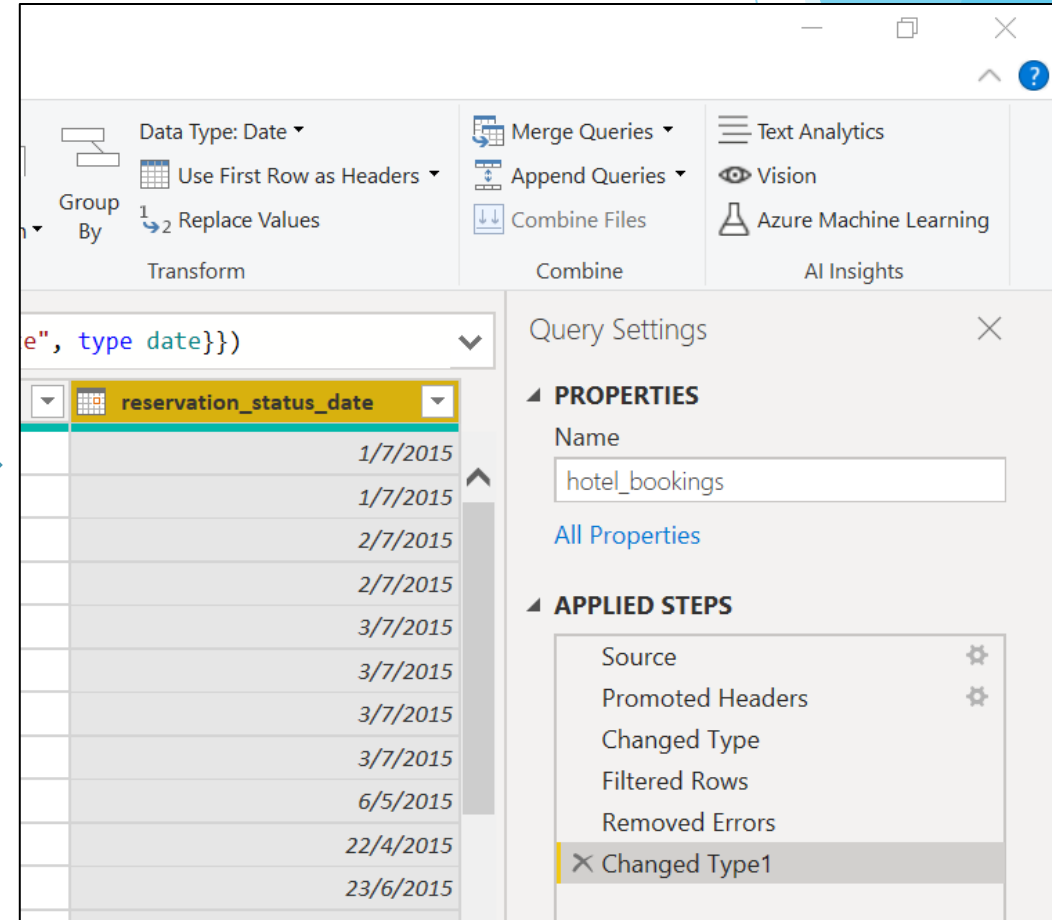
Data Preparations:

► Changed Data Type



The screenshot shows the Power Query Editor interface. The 'Data Type' dropdown is set to 'Text'. The 'reservation_status_date' column is selected in the table. The 'Query Settings' pane on the right shows the 'APPLIED STEPS' list, which includes 'Source', 'Promoted Headers', 'Changed Type', 'Filtered Rows', and 'Removed Errors'. The 'Changed Type' step is highlighted.

reservation_status_date
2015-07-01
2015-07-01
2015-07-02
2015-07-02
2015-07-03
2015-07-03
2015-07-03
2015-07-03
2015-05-06
2015-04-22
2015-06-23
2015-07-05

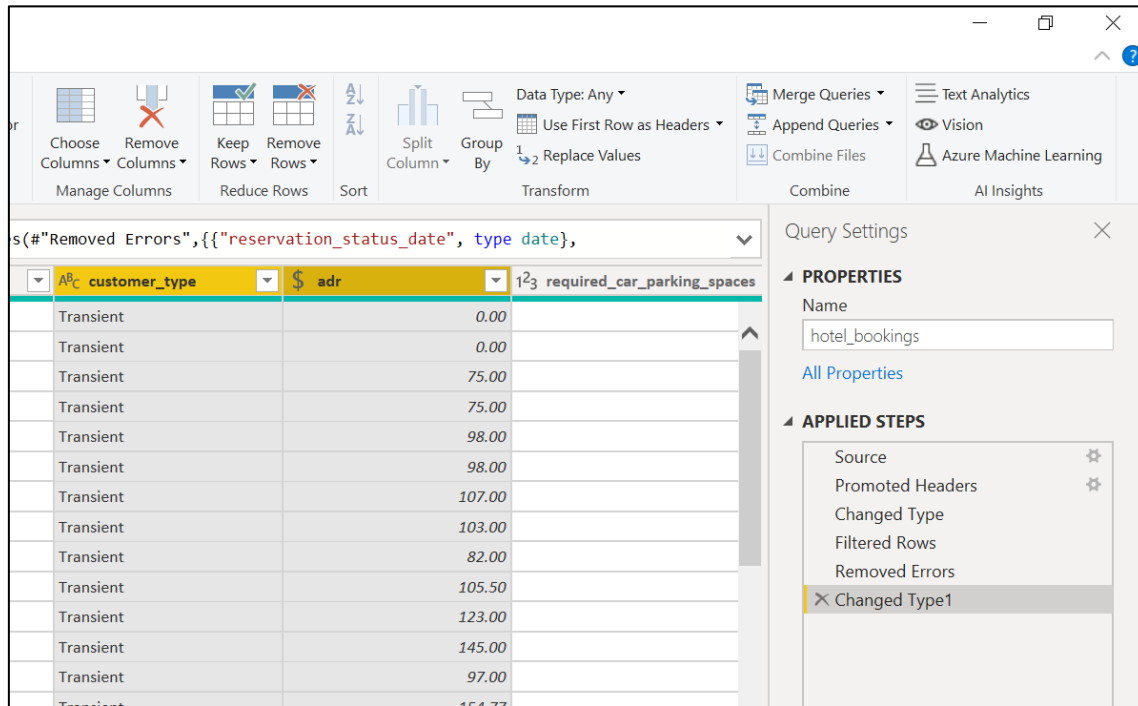


The screenshot shows the Power Query Editor interface after the 'Changed Type' step has been applied. The 'Data Type' dropdown is now set to 'Date'. The 'reservation_status_date' column is selected in the table. The 'Query Settings' pane on the right shows the 'APPLIED STEPS' list, which includes 'Source', 'Promoted Headers', 'Changed Type', 'Filtered Rows', 'Removed Errors', and 'Changed Type1'. The 'Changed Type1' step is highlighted.

reservation_status_date
1/7/2015
1/7/2015
2/7/2015
2/7/2015
3/7/2015
3/7/2015
3/7/2015
3/7/2015
3/7/2015
6/5/2015
22/4/2015
23/6/2015

Data Preparations:

▶ Renamed Columns



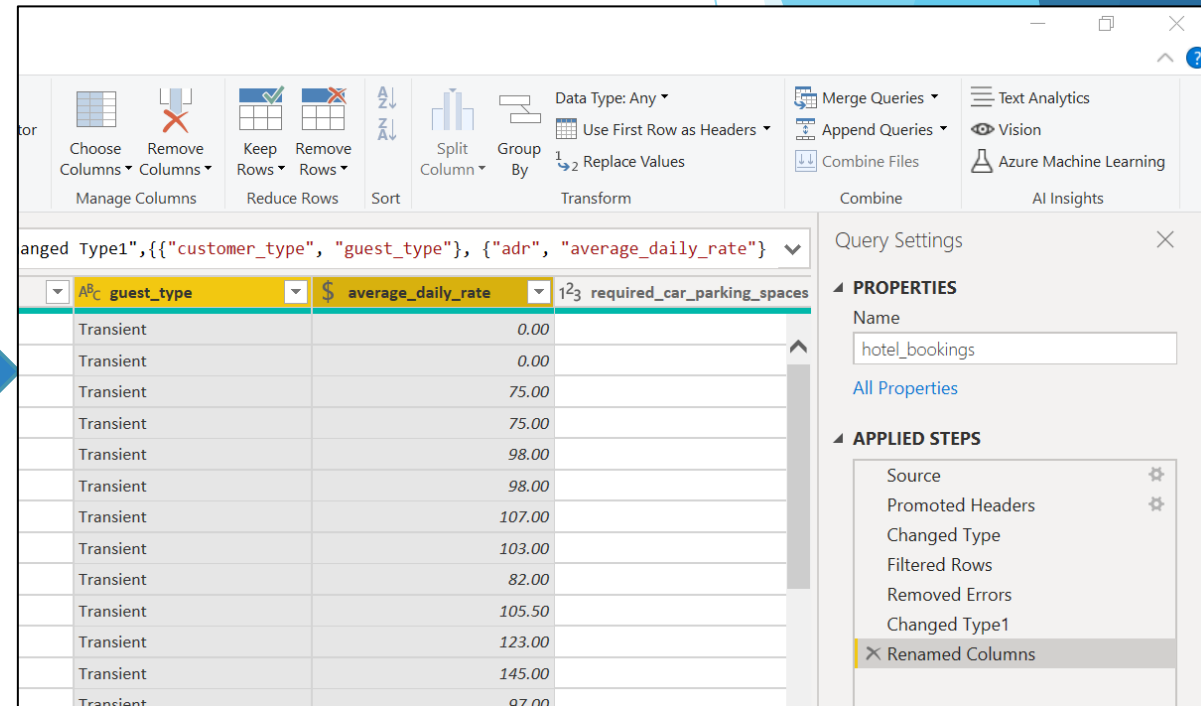
Query Settings

Name: hotel_bookings

APPLIED STEPS

- Source
- Promoted Headers
- Changed Type
- Filtered Rows
- Removed Errors
- Changed Type1

customer_type	adr	required_car_parking_spaces
Transient	0.00	
Transient	0.00	
Transient	75.00	
Transient	75.00	
Transient	98.00	
Transient	98.00	
Transient	107.00	
Transient	103.00	
Transient	82.00	
Transient	105.50	
Transient	123.00	
Transient	145.00	
Transient	97.00	
Transient	154.33	



Query Settings

Name: hotel_bookings

APPLIED STEPS

- Source
- Promoted Headers
- Changed Type
- Filtered Rows
- Removed Errors
- Changed Type1
- Renamed Columns

guest_type	average_daily_rate	required_car_parking_spaces
Transient	0.00	
Transient	0.00	
Transient	75.00	
Transient	75.00	
Transient	98.00	
Transient	98.00	
Transient	107.00	
Transient	103.00	
Transient	82.00	
Transient	105.50	
Transient	123.00	
Transient	145.00	
Transient	97.00	
Transient	154.33	

Data Preparations:

Merged Columns

Query Settings

Name: hotel_bookings

APPLIED STEPS: Source, Promoted Headers, Changed Type, Filtered Rows, Removed Errors, Changed Type1, Renamed Columns, Changed Type2, Added Index, Reordered Columns, **Renamed Columns1**

lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number
0	342	2015 July	
0	737	2015 July	
0	7	2015 July	
0	13	2015 July	
0	14	2015 July	
0	14	2015 July	
0	0	2015 July	
0	9	2015 July	
1	85	2015 July	
1	75	2015 July	
1	23	2015 July	
0	35	2015 July	
0	68	2015 July	
0	18	2015 July	
0	37	2015 July	
0	68	2015 July	



Changed data type to match column data type

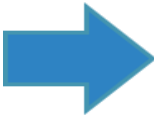
Query Settings

Name: hotel_bookings

APPLIED STEPS: Source, Promoted Headers, Changed Type, Filtered Rows, Removed Errors, Changed Type1, Renamed Columns, Changed Type2, Added Index, Reordered Columns, Renamed Columns1, **Changed Type3**

lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number
0	342	2015	July
0	737	2015	July
0	7	2015	July
0	13	2015	July
0	14	2015	July
0	14	2015	July
0	0	2015	July
0	9	2015	July
1	85	2015	July
1	75	2015	July
1	23	2015	July
0	35	2015	July
0	68	2015	July
0	18	2015	July
0	37	2015	July
0	68	2015	July

Merging



Query Settings

Name: hotel_bookings

APPLIED STEPS: Source, Promoted Headers, Changed Type, Filtered Rows, Removed Errors, Changed Type1, Renamed Columns, Changed Type2, Added Index, Reordered Columns, Renamed Columns1, Changed Type3, Added Custom, **Reordered Columns1**

arrival_date_month	arrival_date_month&year	arrival_date_week_number	arrival_date_day
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	
July	July2015	27	



Final Merged Columns

Query Settings

Name: hotel_bookings

APPLIED STEPS: Source, Promoted Headers, Changed Type, Filtered Rows, Removed Errors, Changed Type1, Renamed Columns, Changed Type2, Added Index, Reordered Columns, Renamed Columns1, **Changed Type3**

Custom Column

New column name: arrival_date_month&year

Custom column formula: = ([arrival_date_month]&[arrival_date_year])

Available columns: Booking_id, hotel, is_canceled, lead_time, arrival_date_year, arrival_date_month, arrival_date_week_number, arrival_date_day_of_month

Learn about Power Query formulas

No syntax errors have been detected.

OK Cancel

Hotel Location from Dataset:



► Portugal: Lisbon (City Hotel)



Hotel Location from Dataset:



► Portugal: Algarve (Resort Hotel)



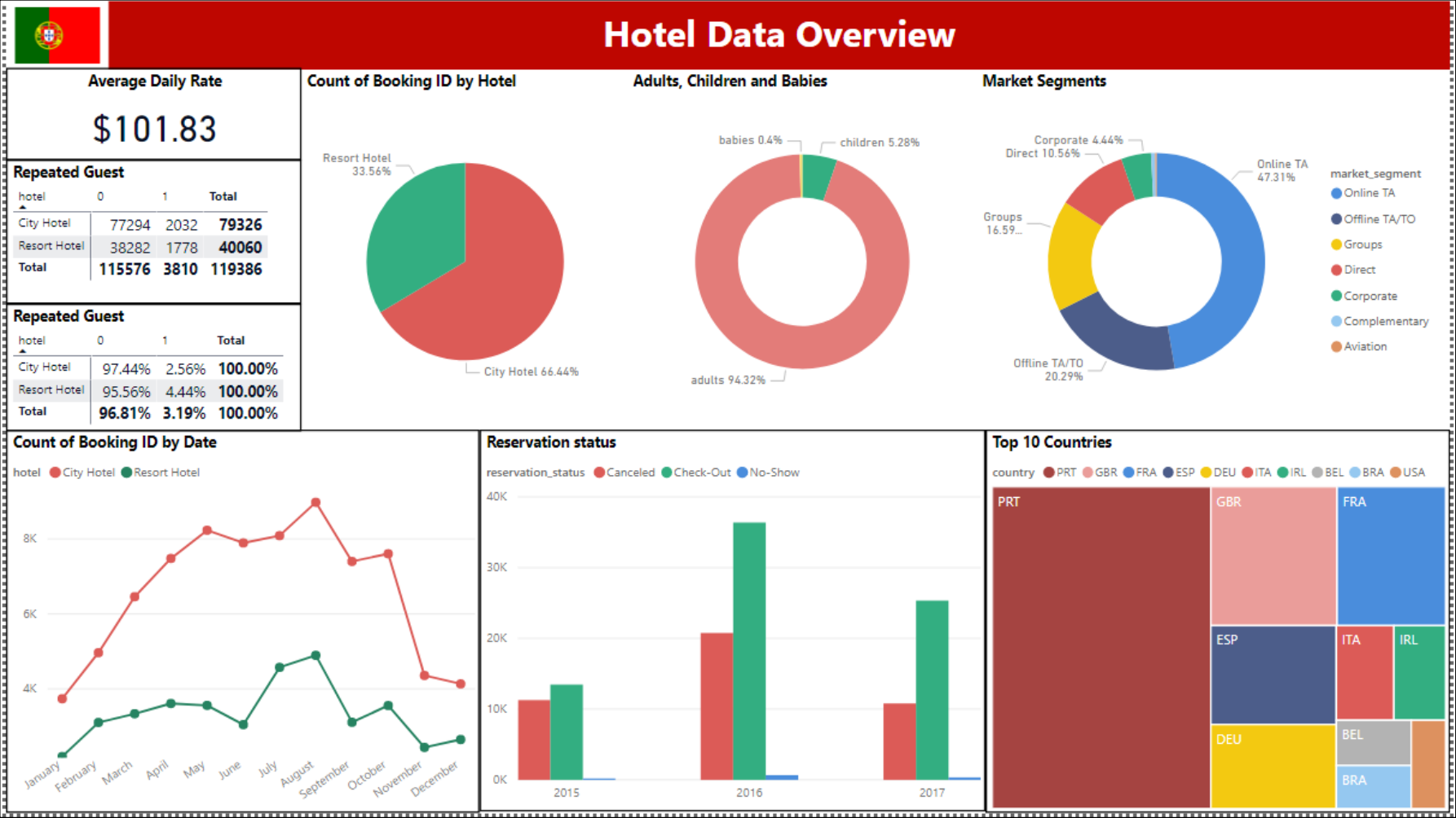
Hotel Stakeholders:

- ▶ Hotel General Managers
- ▶ Hotel Owners
- ▶ Hotel Regional Managers (Hotel Chains)
- ▶ Investors

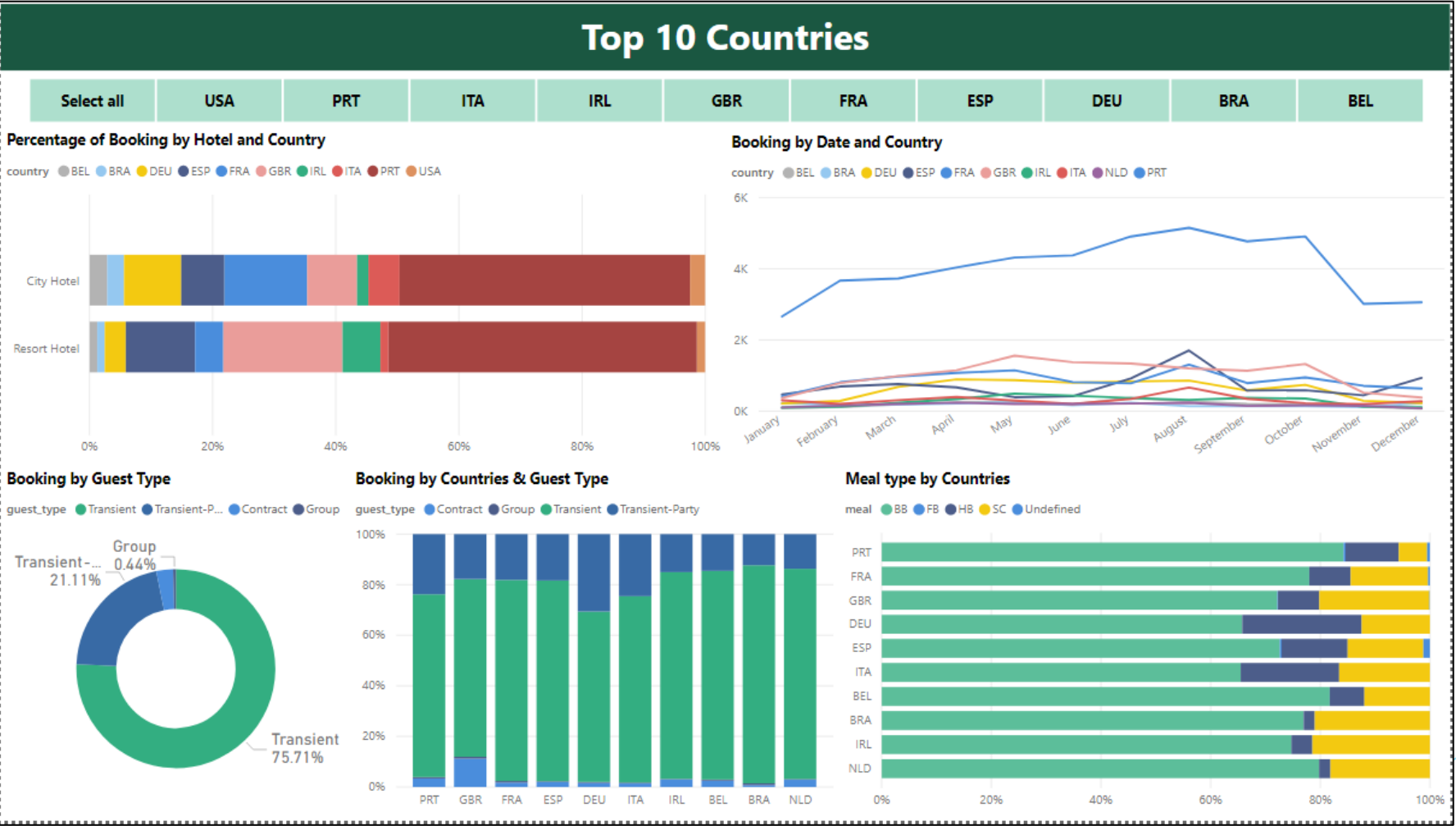
Business Questions:

- ▶ Any visible trends in arrival and occupancy?
- ▶ Which is the biggest source of guest for the hotel?
- ▶ Which room type brings the largest revenue?
- ▶ Any potential area to expand business in?

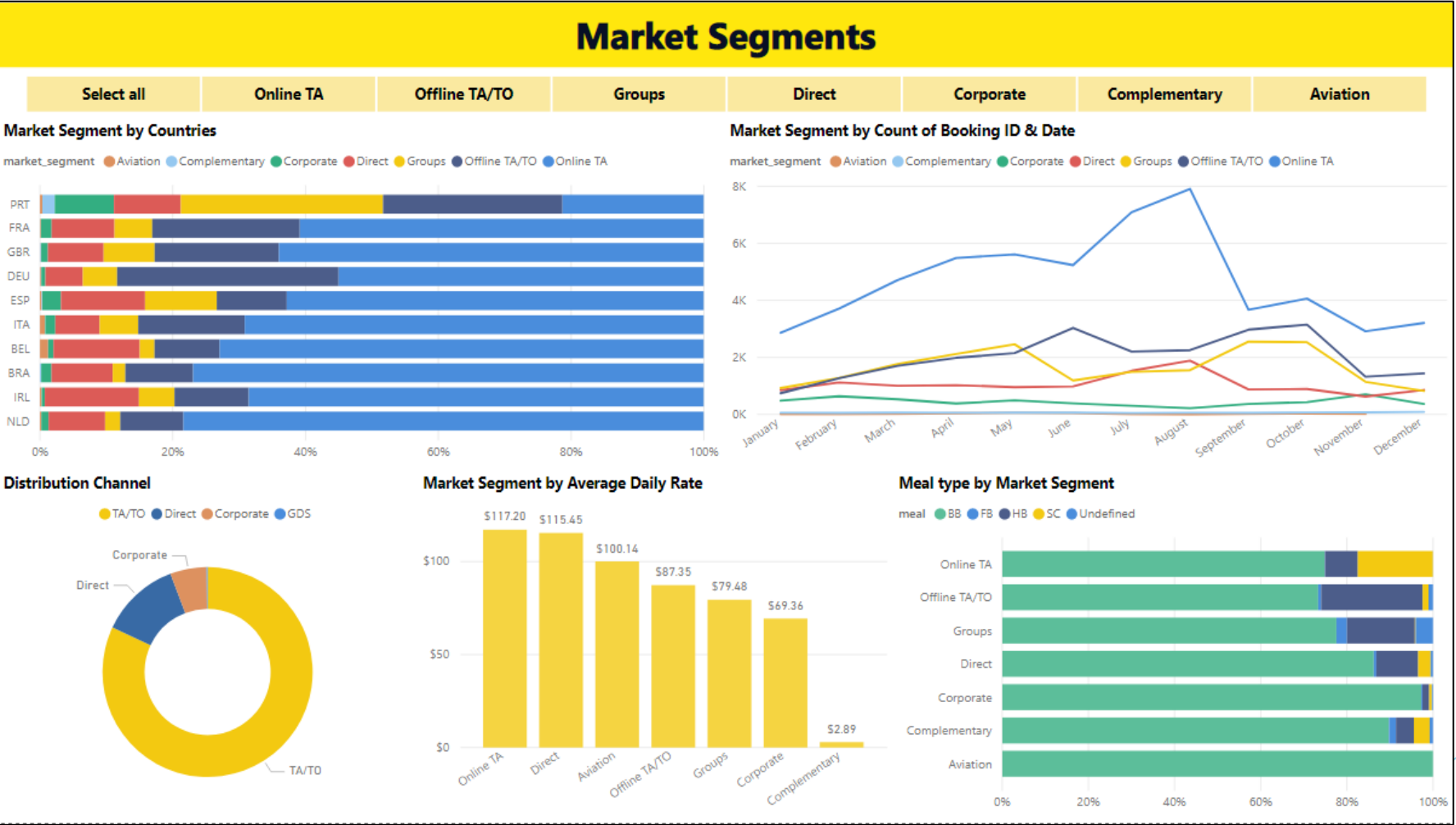
Dashboard 1:



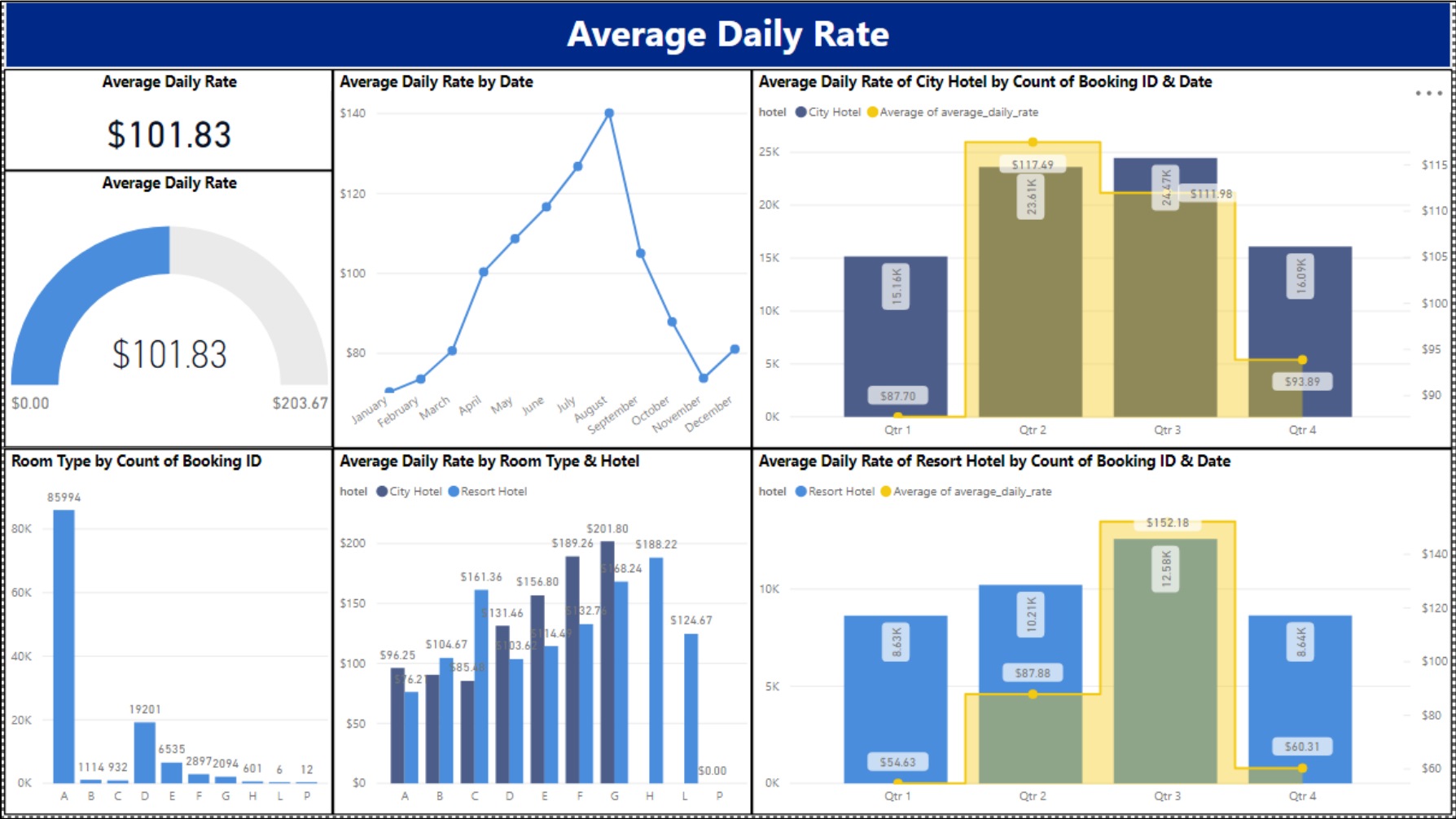
Dashboard 2:



Dashboard 3:



Dashboard 4:



Conclusion (Insights):

- ▶ There is seasonality for both resort and city hotels
- ▶ Biggest hotel guest nationality is the Portuguese
- ▶ Almost half of the total number of guest uses online Travel Agency
- ▶ Average room rate varies from room type