



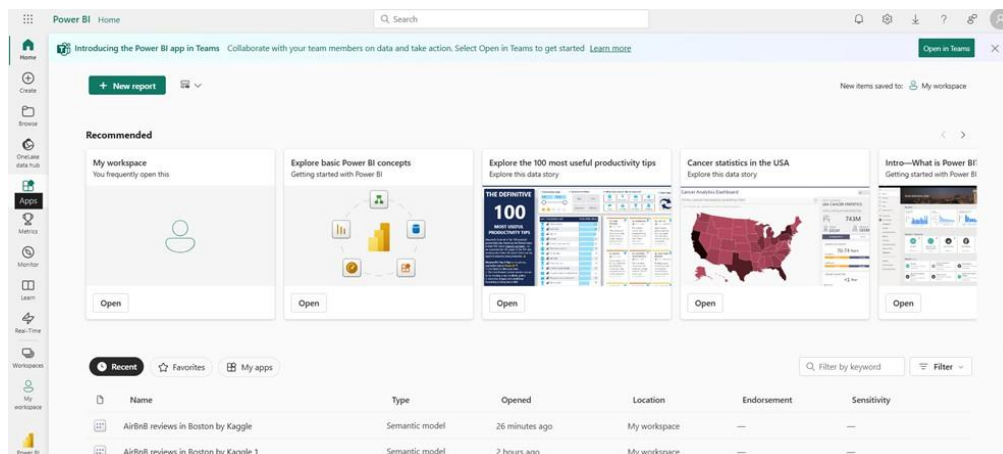
**Data Science and
Database Technology**

Politecnico di Torino

PowerBI -Practice 2

1. Login

Connect to [PowerBI](#), login with your Microsoft Account (the one linked to you polito id or your personal one). Use the **online** version of PowerBI.



2. New report on Airbnb Boston reviews

To create a new report from scratch, a data source must be identified. To this aim, a portion of the <https://www.kaggle.com/datasets/airbnb/boston> has been uploaded into a shared Google Sheets to be used as data source for PowerBI.

- the Google Sheets, with approximately 10k reviews to be used as data source, is available at <https://docs.google.com/spreadsheets/d/1a2c9vCMFFfDXmhjoEoX2EwS2lYTbqE4WfZY72TXW9co/edit?gid=285360760#gid=285360760>

Create a new report

• After uploading your data successfully, you have to make sure that you are signed in to PowerBI using **your institutional credentials**.

- Click on "Next", and now choose the appropriate sheet (Here, **Reviews Query DW**)
- Click on "Create"

Get data
Choose data

Search

Display options

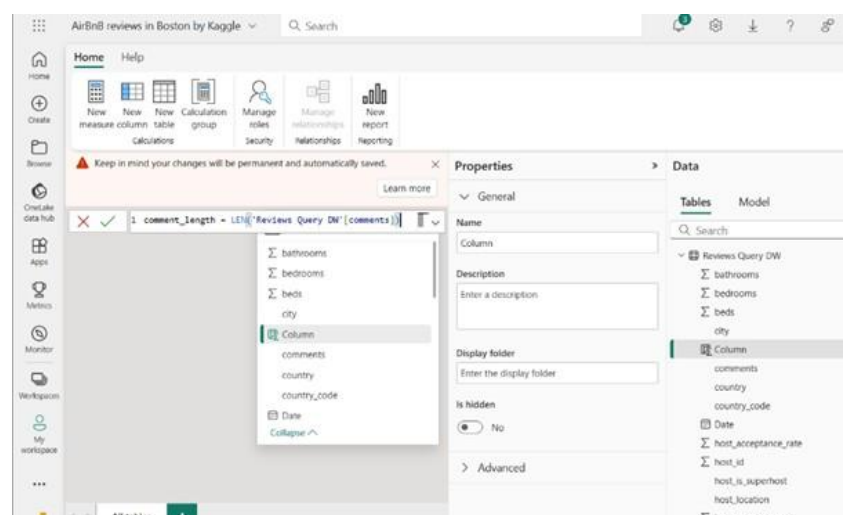
- Excel workbook [3]
 - Listings
 - Reviews Query DW
 - Sheet1
- Suggested tables [1]
 - Table 1 (Listings)

Reviews Query DW

listing_id	id	Date	reviewer_id	reviewer_name	comments
9857	92488773	8/8/2016	56470796	Danielle	Seamless was easy to work with -
9857	90834004	8/2/2016	86373945	Courtney	Seamless was really great about!
9857	81205609	6/21/2016	74520248	Patrick	We had a fantastic stay in Bostc
9857	52407219	10/29/2015	42949625	Heather	Excellent location, good sized h
9857	51958656	10/25/2015	45617313	John	The apartment was even better
9857	46422364	9/10/2015	2642405	Mike	Awesome place, nicer than the i
9857	37160926	7/5/2015	11002414	Stephanie	The name seamless says it all. T
9857	33579553	5/30/2015	15020381	Kristina	Everything worked perfect, from
9857	32775740	5/21/2015	18106873	Chris	Great location and great place t
9857	31603560	5/7/2015	2090801	Matt	Great apartment. Not huge but
9857	14951136	6/30/2014	11434319	David	Apartment is small, but was wel
9857	7978534	10/11/2013	4257452	Maya	Excellent apartment. Very clean,
9857	7663739	9/29/2013	5804393	Harry	Brilliant location in proximity to
9857	7362736	9/17/2013	3312378	Don	We had a great stay and enjoye
					I would stay there again but fut
9857	7303002	9/15/2013	4062028	Stephanie	Great apartment with plenty of
9857	5393746	6/28/2013	5753987	David	This 2BR was is a fantastic locat
9857	5258160	6/21/2013	5881085	Jennifer	Our experience renting and stay

Dimensions, metrics, and transformations

- Check the **type** of each field. Measures and dimensions are interpreted **automatically**. Although, you can safely go on if you think that they have been misinterpreted.
- **Task:** Go to **Modeling** and create a new columns where it contains the length of the comments (count the number of chars of the comment field). Note that you can use the following function **LEN()**



Analyse the data

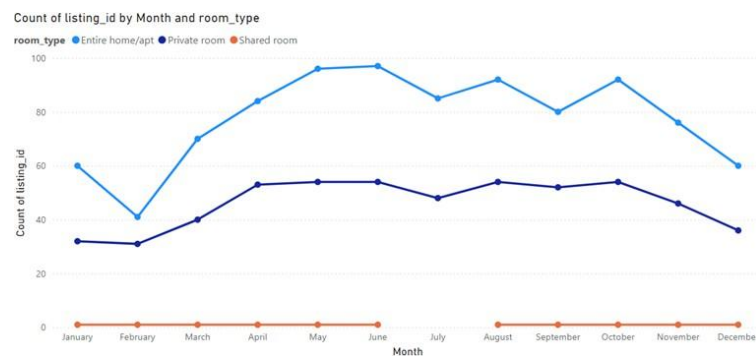
Analyse the data by building the following visualizations. Then, explore and create new visualizations to find interesting insights on your own.

- **Analysis (1):** compare the trend of the **average** (pay attention to that) length of the review "comments" (number of chars) vs the average "review_scores_rating" for different "proptert_type". Sort the data by

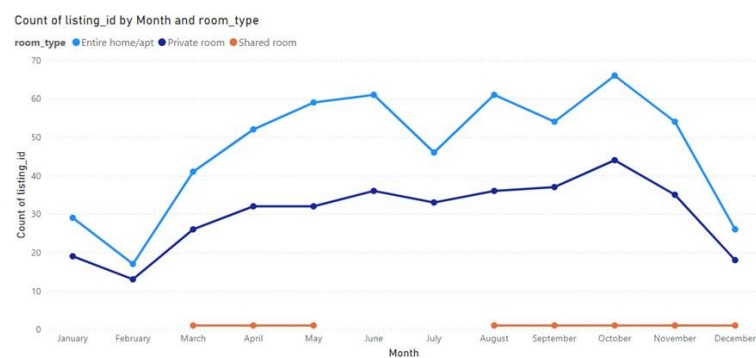
descending average length of comments. Allow end-users to filter the data under analysis by selecting a date range of their choice.



- **Analysis (2):** compare the trend of the number of different "listing_id" reviewed, for each "room_type", and for each month of the year. Allow end-users to filter the data under analysis by selecting a date range and the type of superhost (true/false).

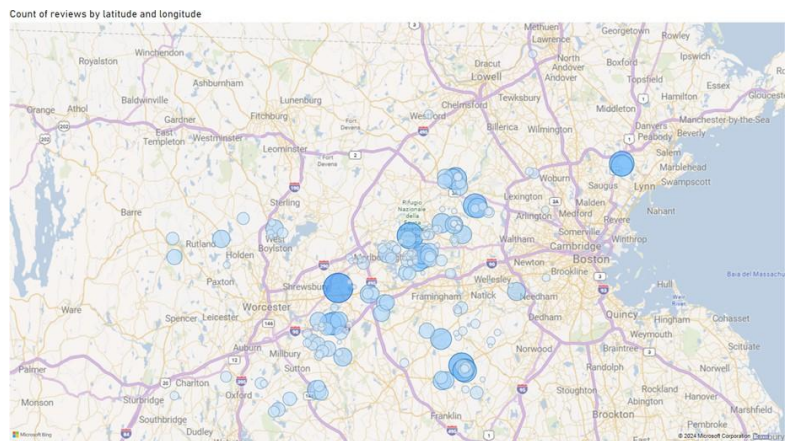


For example, if you filter data to contain only data related to year 2011 and 2013, you get the following graph.



- **Analysis (3):** analyse the number of different reviewers for each location (lat, long).

Note that the Kaggle dataset of the Airbnb reviews is in Boston, **Massachusetts**, US



- **Analysis (4):** Visualize, for each property type and for each year, the average rating score values sorted by ascending property type and by descending mean rating_score_value. Exclude possible null values for the attribute property_type.

property_type	Year	Average of review_scores_rating
Apartment	2009	93.63
Apartment	2010	92.84
Apartment	2011	91.27
Apartment	2012	91.27
Apartment	2013	91.23
Apartment	2014	91.20
Apartment	2015	91.22
Apartment	2016	92.06
Bed & Breakfast	2011	87.19
Bed & Breakfast	2012	87.63
Bed & Breakfast	2013	87.38
Bed & Breakfast	2014	87.61
Bed & Breakfast	2015	86.56
Bed & Breakfast	2016	87.36
Boat	2010	95.00
Boat	2011	94.75
Boat	2012	94.71
Boat	2013	94.75
Boat	2014	94.33
Boat	2015	94.88
Boat	2016	94.80
Condominium	2011	91.00
Condominium	2013	92.00
Condominium	2014	93.00
Condominium	2015	95.57
Condominium	2016	94.63
House	2009	91.00
House	2010	91.82
Total		91.48

- **Analysis (5):** Visualize, for each year and for each room type, the total count of top-scored reviews (review_score_value > 10).

Compare the obtained results with the count of the distinct listing_id reviewed.

