

Hello! As a fellow data analyst, this is a fantastic project. Analyzing the real-world impact of regulations is a high-value task. Based on the data dictionaries you provided and enriched with external information, here is a breakdown of the columns and their intended analytical purpose, especially in the context of your Paris project.

Listings Data Dictionary (Listings_data_dictionary.csv)

First, here is the dictionary you provided for the listings data:

Field	Description
listing_id	Listing ID
name	Listing Name
host_id	Host ID
host_since	Date the Host joined Airbnb
host_location	Location where the Host is based
host_response_time	Estimate of how long the Host takes to respond
host_response_rate	Percentage of times the Host responds
host_acceptance_rate	Percentage of times the Host accepts a booking request
host_is_superhost	Binary field to determine if the Host is a Superhost
host_total_listings_count	Total listings the Host has in Airbnb
host_has_profile_pic	Binary field to determine if the Host has a profile picture
host_identity_verified	Binary field to determine if the Host has a verified identity
neighbourhood	Neighborhood the Listing is in
district	District the Listing is in
city	City the Listing is in
latitude	Listing's latitude

longitude	Listing's longitude
property_type	Type of property for the Listing
room_type	Type of room type in Airbnb for the Listing
accommodates	Guests the Listing accomodates
bedrooms	Bedrooms in the Listing
amenities	Amenities the Listing includes
price	Listing price (in each country's currency)
minimum_nights	Minimum nights per booking
maximum_nights	Maximum nights per booking
review_scores_rating	Listing's overall rating (out of 100)
review_scores_accuracy	Listing's accuracy score based on what's promoted in Airbnb (out of 10)
review_scores_cleanliness	Listing's cleanliness score (out of 10)
review_scores_checkin	Listing's check-in experience score (out of 10)
review_scores_communication	Listing's communication with the Host score (out of 10)
review_scores_location	Listing's location score within the city (out of 10)
review_scores_value	Listing's value score relative to its price (out of 10)
instant_bookable	Binary field to determine if the Listing can be booked instantly

Analytical Purpose of Key Listing Columns

Here is the breakdown of what these columns are used for in analysis:

1. Key Identifiers & Relationships

- listing_id**: This is the **Primary Key (PK)** for this dataset. Every row represents a unique property.

1. Waleed Mouhammed

12 November 2025 at 09:15:26

It seems that the long-term rental is the main concern for the regulator as he should ensure that there is enough units for long-term rental for the people inside the city!

- **host_id**: This is the **Foreign Key (FK)** that links listings to a specific host. This is crucial as it creates a **one-to-many relationship** (one host can have many listings).
 - **Analytical Use**: You can **GROUP BY host_id** to analyze host-level behavior. For your project, this is essential for identifying **professional hosts** (e.g., **host_total_listings_count > 1**) versus casual hosts. **Regulators are often most concerned with professional hosts who remove multiple properties from the long-term rental market.**

2. Host-Level Features

- **host_is_superhost**: A binary (True/False) flag. This is a key feature for **segmentation**.
 - **Analytical Use**: You can analyze if the new regulations disproportionately affected Superhosts versus regular hosts.
- **host_response_time**, **host_response_rate**: These are host quality metrics. You can use them as features in a model to see if host engagement correlates with price or review scores.

3. Geospatial Features

- **latitude**, **longitude**: These are for geospatial analysis.
 - **Analytical Use**: You can use these to plot listings on a map of Paris (e.g., a heatmap) to see if regulations caused a drop in listings in specific, high-demand *arrondissements* (districts).
- **neighbourhood**, **district**: Categorical variables for grouping by area.

4. Critical Attributes for Regulatory Analysis

- **property_type** (e.g., 'Apartment', 'House', 'Loft')
- **room_type** (e.g., 'Entire home/apt', 'Private room', 'Shared room')
 - **Critical Distinction**: My search confirms that this is the *most important* distinction for your project. **property_type** is the physical building, but **room_type** is what's being rented.
 - **Analytical Use**: Parisian regulations heavily target **'Entire home/apt'** listings, as these are often "secondary residences" rented full-time to tourists. **'Private room'** listings (where the host lives on-site in their "primary residence") are often exempt from the strictest rules, such as Paris's 120-day-per-year cap. Your primary analysis will likely involve segmenting by this column.

5. Policy & Pricing Variables

- **price**: This is your primary **dependent variable** (outcome variable).
 - **Analytical Use**: Did the average price of available listings *increase* after regulations (due to reduced supply) or *decrease* (due to lower demand or other factors)?
- **minimum_nights**: This is an *extremely* important column for your project.
 - **Analytical Use**: Short-term rental laws (like in Paris or NYC) are often based on stay duration. My research indicates many hosts in Paris have set their

`minimum_nights` to **30 or more**. This is a common strategy to **bypass short-term regulations**, reclassifying the listing as a "mid-term rental" or "mobility lease" (*bail mobilité*), which are not subject to the same caps. You must track the distribution of this variable before and after the new rules.

6. Performance Metrics (KPIs)

- `review_scores_rating`: The main Key Performance Indicator (KPI) for guest satisfaction (out of 100).
- `review_scores_...` (**value, cleanliness, etc.**): These are sub-metrics.
 - **Analytical Use**: You can measure if the *quality* of listings changed. For example, did the `review_scores_value` drop after regulations, suggesting guests felt they were paying more for less?

Reviews Data Dictionary (`Reviews_data_dictionary.csv`)

Here is the dictionary you provided for the reviews data:

Field	Description
<code>listing_id</code>	Listing ID
<code>review_id</code>	Review ID
<code>date</code>	Review date
<code>reviewer_id</code>	Reviewer ID

Analytical Purpose of Review Columns

This file is your **transactional data**, or "event log." Its primary purpose is to link listings to activity over time.

- `review_id`: The **Primary Key (PK)** for this table. Each row is a unique review.
- `listing_id`: The **Foreign Key (FK)** used to **join** this table to the `Listings` table.
- `reviewer_id`: An identifier for the guest. You could use this to analyze reviewer behavior (e.g., how many reviews a single reviewer leaves), but it's often anonymized.

- **date**: This is arguably the **most important column in the entire dataset** for your specific project.
 - **Analytical Use**: This is your time-series variable. By aggregating this column (e.g., by month), you can perform the **pre-and-post-regulatory analysis**. You can count the number of reviews per month, or calculate the average `review_scores_rating` per month, and plot this time-series. You can then add a vertical line on the date the regulations were implemented to visually (and statistically) assess their impact.

Summary

In short, the **Listings** file gives you the "what" and "where" (the features of the properties), while the **Reviews** file gives you the "when" (the time-series data), which is essential for your impact assessment.

Good luck with your analysis!

2. Waleed Mouhammed
12 November 2025 at 09:26:37
For travel apps such as Airbnb and Bookings, it's all about short-term rental. So, almost all the regulations goes around the short

3. Waleed Mouhammed
12 November 2025 at 09:32:24
This was the first step; is to identify the primary and secondary residence. If you stay at your residence for 8 months at least per year then this is your primary residence, else it will be secondary residence.

4. Waleed Mouhammed
12 November 2025 at 09:34:17
Here is the thing: you can rent a private room in your entire primary residence without being capped, but if you are going to rent your entire residence then you are capped to 120 days per year which is 4 month - remember that your primary residence is primary when and ONLY when you stay at it for 8 months per year.

Of course. This is a critical part of your project. Understanding the "why" and "when" of these regulations will be the key to structuring your pre-and-post-analysis.

Here is a timeline of the key regulations, a simple explanation of what they are, the regulator's perspective, and the likely effects on the market.

2 Paris Short-Term Rental Regulation: A Timeline for Analysis

This timeline outlines the major interventions. For your project, the **late 2017 / early 2018** period is likely the most significant "before and after" moment to monitor.

Date	Regulation / Event	Simple Explanation
2014	ALUR Law	This law set the foundation. It officially defined a "primary residence" as the place you live for at least 8 months per year. Anything else is a "secondary residence." This distinction is the single most important concept in Parisian regulation.
2016-2017	Digital Republic Law	This is the big one. This law introduced the 120-day-per-year cap . An owner is only allowed to rent their <i>entire primary residence</i> for a maximum of 120 days per year. Renting a <i>private room</i> in your primary residence is not capped.
Oct 1, 2017	Registration Requirement	Paris mandated that all hosts renting their primary residence (even for one night) must declare it with the city.

5. Waleed Mouhammed
12 November 2025 at 09:38:00
Here we are talking about the entire residence NOT the private

6. Waleed Mouhammed
12 November 2025 at 09:39:23
This is the regulator perspective, protecting their citizens from the housing shortage which consequently leads to higher prices.

Dec 1, 2017	Registration B	This is a critical date for your data . From this point, all short-term rental listings were legally required to display a 13-digit registration number (e.g., 7510101234567). This rule applies to all platforms, including Airbnb and Booking.com.
Jan 1, 2018	Automated Ca	Airbnb began automatically enforcing the 120-day cap for listings in the 4 central districts (arrondissements 1-4), automatically blocking a host's calendar after they reached the limit. This was later expanded.
2018	ELAN Law	This law gave cities more power. It introduced heavy fines for both hosts and platforms. Platforms (like Airbnb) could be fined thousands of euros <i>per illegal listing</i> they failed to remove.
2020	EU Court of Ju	A major legal victory for Paris. The court confirmed that the city's rules (registration, 120-day cap, etc.) were justified by "a compelling reason of general interest"—specifically, the housing shortage .

7. Waleed Mouhammed
12 November 2025 at 09:41:14
Here we are! It becomes 90 days!

8. Waleed Mouhammed
12 November 2025 at 09:42:51
That's it! This is the core reason, protecting the Parisians.

9. Waleed Mouhammed
12 November 2025 at 09:44:14
So, the hosts who have multiple listings will be targeted heavily in our analysis to check how were they impacted due to these regulations.

Nov 2024	Loi "Le Meur"	A new national law was passed to give local mayors even more tools to regulate short-term rentals, including enhanced control over "change of use" and new tax policies.	7
Jan 1, 2025	Cap Reduced to 90 days	Using the new "Le Meur" law, the City of Paris voted to reduce the annual cap for primary residences from 120 days down to 90 days . This is the most recent major change and provides a second "event" for your analysis.	

The Regulator's Point of View (The "Why")

City regulators, led by the Paris City Hall, implemented these rules to solve several specific problems:

- 8
- 9
1. **To Combat the Housing Crisis:** This is the #1 reason. Regulators believe that professional investors were buying apartments meant for residents and turning them into full-time, "disguised" hotels. **This removes housing from the long-term rental market, which in turn drives up rents and property prices for Parisians.**
2. **To Preserve Neighborhoods:** They want to prevent residential areas from becoming "tourist-only" zones, which can lead to neighborhood nuisances (noise, parties) and the loss of local shops (like bakeries) being replaced by souvenir stores.
3. **To Ensure Fair Competition:** The traditional hotel industry must follow strict rules on safety, labor, and taxes. Regulators see the old, unregulated Airbnb market as creating an unfair, untaxed, and unsafe "shadow" hotel industry.
4. **To Distinguish "Sharing" from "Business":** The city's official stance is that they are *not* against residents who "share" their own home for a few weeks a year (the original idea of Airbnb). **They are against professional operators running a full-time, multi-property rental business in residential buildings.**

How These Changes Affect Platforms (Airbnb, Booking.com)

10. Waleed Mouhammed

12 November 2025 at 09:45:45

They will be penalized if they do not!

11. Waleed Mouhammed

12 November 2025 at 09:53:18

The main idea is to remove the secondary residence from the whole process.

These regulations fundamentally changed the business environment for platforms like Airbnb and Booking.com.

1. **Increased Responsibility & Compliance:** The platforms are no longer just "marketplaces." They are now legally responsible for enforcing the rules. They must:
 - Verify and display the 13-digit registration number.
 - Automatically block calendars for primary residences that hit the 90/120-day cap.
 - Share data with city and tax authorities.
 - **Remove illegal listings or face massive fines.**
2. **A Squeeze on "Entire Home" Supply:** The rules are *designed* to remove a specific type of listing: **the "Entire home/apt" listing that is a secondary residence**. This is the most profitable type of listing, so you should expect to see a significant impact on this category in your dataset.
3. **A Predictable Shift in Listing Strategy (What to look for in your data):**
 - **Decrease in "Entire Home" listings:** As illegal listings are removed.
 - **Increase in "Private Room" listings:** This `room_type` is not subject to the 90/120-day cap, making it a safer, more stable option for hosts.
 - **A Jump in `minimum_nights`:** This is a crucial analytical point. Hosts will try to bypass these *short-term* regulations by setting their `minimum_nights` to 30 or more. This reclassifies the rental as a "mid-term" or "mobility lease" (*bail mobilité*), which is not subject to the cap. You should absolutely plot the distribution of `minimum_nights` over time.

That's a critical point to clarify, and it's actually the other way around: the regulations for **secondary residences** are the *most* strict, to the point of being a near-total ban.

This is the core of the entire regulatory policy. Here's the simple breakdown:

1. Primary Residence (The "Exception")

- **What it is:** The home you live in for at least 8 months per year.
- **The Rule:** You are given a *special allowance* to rent this property on a short-term basis, but **only up to a cap** (which was 120 days, and is now 90 days). You must also have a registration number.
- **The Logic:** The regulator is "tolerating" home sharing. They are okay with you renting your *own* home while you're on vacation, but not so much that your home becomes a full-time hotel.

2. Secondary Residence (The "Target")

- **What it is:** Any property that is *not* your primary residence (e.g., an investment property, a second home).
- **The Rule:** It is **illegal by default** to rent a secondary residence on a short-term basis (even for one night).
- **The Only Legal Way:** To do it legally, an owner must:
 1. Apply for a "**change of use**" authorization from the city.
 2. This officially converts the property's legal status from "residential" to "commercial."
 3. To get this, the owner must provide "**compensation.**" This usually means buying *another* commercial property (like an office or shop) of an equivalent or larger size in the same area and permanently converting it *into* a new residential apartment.

Summary

The 90/120-day cap isn't the main rule; it's the *exception* for primary homes. The **main rule is a ban on short-term rentals in secondary residences.**

The "compensation" mechanism is so expensive and difficult that it effectively stops investors from buying apartments to rent on Airbnb. This is precisely what the regulator wants.

So, in your analysis, when you see a listing for an "Entire home/apt," the key question is whether it's a primary residence (capped at 90/120 days) or an illegal secondary residence.

You're right to focus on this; it's the most clever and important data signal you'll find.

Here is the concept broken down:

1. The Regulations Target *Only* "Short-Term" Stays

Think of the regulations (like the 90/120-day cap) as a box. This box *only* applies to rentals that are legally defined as "**short-term rentals**."

2. The Law Defines "Short-Term" by Duration

In France, as in many countries, the law draws a line between different types of rentals. A stay for a few nights (a tourist stay) is not treated the same as a 6-month lease.

- **Short-Term Rental (STR):** This is what the caps target. It's the typical tourist model: renting for a few nights or a week.
- **Mid-Term / "Mobility Lease" (*bail mobilité*):** This is a different legal category. In France, this lease was specifically created for students, business travelers, or people in transition who need a furnished place for **one to ten months**.
- **Long-Term Rental:** This is a standard residential lease, typically one year or more.

3. `minimum_nights` is the "Switch"

The `minimum_nights` setting in your dataset is the *host's* way of deciding which legal "box" they want to be in.

- **If a Host sets `minimum_nights = 2`:**
 - They are clearly operating as a **Short-Term Rental**.
 - They are *inside* the regulation box.
 - They *must* follow the 90/120-day cap.
- **If a Host sets `minimum_nights = 30`:**
 - They are *filtering out* all short-term tourists. They are *only* accepting guests who will stay for at least a month.
 - By doing this, they can legally reclassify their operation. They are no longer a Short-Term Rental; they are now a **Mid-Term Rental** (e.g., offering a *bail mobilité*).

- Because they are *outside* the short-term regulation box, **the 90/120-day cap no longer applies to them.**

Why Hosts Do This: The "Workaround"

This is a direct, strategic response to the regulations.

Regulator's Goal: "We want to stop you from renting your apartment out every weekend of the year to tourists."

Host's Response: "Fine. I will stop renting to weekend tourists. Instead, I will set my `minimum_nights` to 30. Now I will rent to a student for 3 months, a remote worker for 2 months, and a business consultant for 4 months. I am still renting it 'short-term' in spirit, but legally, I am in a different category that is not capped."

What You Should Look For (Your Analytical Plan)

You should plot a histogram or bar chart of the `minimum_nights` column.

1. **Before the Regulations (e.g., 2016):** You will see almost all listings clustered at `minimum_nights = 1, 2, or 3`.
2. **After the Regulations (e.g., 2019):** You will still see that cluster, but you will very likely see a **new, large spike** suddenly appear in your chart at or around `minimum_nights = 28, 30, or 31`.

This spike represents the hosts who have "jumped" from the regulated short-term market to the unregulated mid-term market. It is a perfect, measurable footprint of the regulations' impact on host behavior.