Low Level Design

New Airbnb Pollution Analysis

Written By	Shubham Tembhurne
Document Version	0.3
Last Revised Date	14/05/25

DOCUMENT CONTROL

Change Record:

VERSION	DATE	AUTHOR	COMMENTS
0.1	02- May - 2025	Shubham temburne	Introduction and architecture defined
0.2	02 - May - 2025	Shubham tembhurne	Architecture & Architecture description appended and updated.

Contents

1. Introduction

1.1 What is Low-Level design document?

The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the Us Pollution Analysis dashboard. LDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

1.2 Scope

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

1.3 Project Introduction

The goal of this project is to analyze the air pollution is the contamination of air due to the presence of substances in the atmosphere that are harmful to the health of humans and other living beings, or cause damage to the climate or to materials.

2.Problem Statement

Given dataset include all information about host, listed properties, geographical location, prices reviews and all other required metrics. Analyze the given dataset make different predictions and draw meaningful conclusion in order to grow the business. Also state what can we learn from different predictions.

2. Dataset Information

Id: Listing ID of the property

State code: Code numbers belong to each state

Name: Name of the listed property.

host_id : ID of the property owner.

host_name : Name of the property owner.

neighbourhood_group : Location at which property located.

Neighbourhood : Area in which property located.

Latitude: Latitude coordinate.

Longitude: Longitude coordinate.

room_type : Type of the room (Entire Home/ Appt, Private Room, Shared Room)

Price: Price in Dollars.

Minimum_nights : Amounts of minimum night stay at property

number_of_reviews : No. Of reviews

last_review : last review on which date.

reviews_per_month : Numbers of reviews per months.

calculated_host_listings_count: Count of properties listed on that host.

availability_365: Number of days when listing is available for booking

3. Architecture

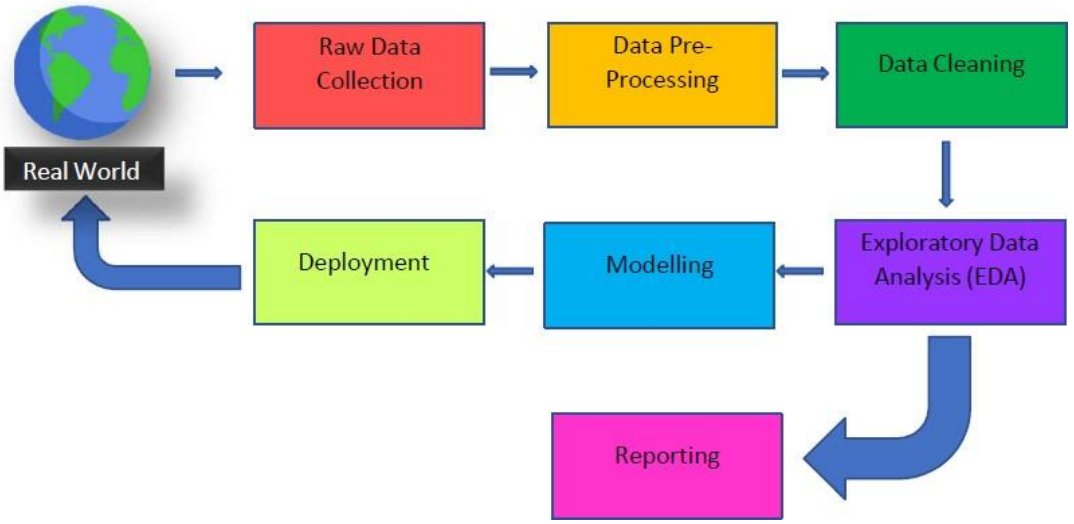


Tableau Server Architecture

Tableau has a highly scalable, n-tier client-server architecture that serves mobile clients, web clients and desktop-installed software. Tableau Server architecture supports fast and flexible deployments.

The following diagram shows Tableau Server's architecture:

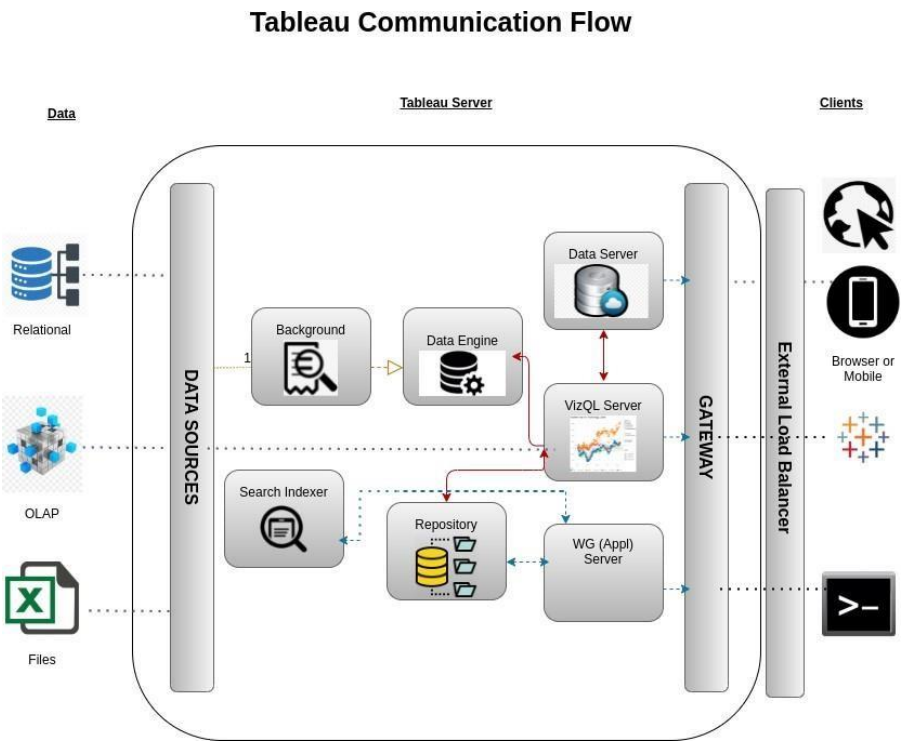


Tableau Server is internally managed by the multiple server processes.

1). Gateway/Load Balancer

It acts as an Entry gate to the Tableau Server and also balances the load to the Server if multiple Processes are configured.

2) Application Server: -

Application Server processes (wgserver.exe) handle browsing and permissions for the Tableau Server web and mobile interfaces. When a user opens a view in a client device, that user starts a session on Tableau Server. This means that an Application Server thread starts and checks the permissions for that user and that view.

3) Repository: -

Tableau Server Repository is a PostgreSQL database that stores server data. This data includes information about Tableau Server users, groups and group assignments, permissions, projects, data sources, and extract metadata and refresh information.

4) VIZQL Server: -

Once a view is opened, the client sends a request to the VizQL process (vizqlserver.exe). The VizQL process then sends queries directly to the data source, returning a result set that is rendered as images and presented to the user. Each VizQL Server has its own cache that can be shared across multiple users

5) Data Engine: -

It Stores data extracts and answers queries.

6) Backgrounder: -

The backgrounder Executes server tasks which includes refreshes scheduled extracts, tasks initiated from tabcmd and manages other background tasks.

7) Data Server: -

Data Server Manages connections to Tableau Server data sources

It also maintains metadata from Tableau Desktop, such as calculations, definitions, and groups.

4.Architecture Description

1. Raw Data Collection

The Dataset was taken from Kaggle Provided.

2. Data Pre-Processing

Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Model performance depends on the quality of data fed to the model to train.
This Process includes)
Handling Null/Missing Values
b) Handling Skewed Data
c) Outliers Detection and Removal

3. Data Cleaning

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.
a) Remove duplicate or irrelevant observations
b) Filter unwanted outliers
c) Renaming required attributes

4. Reporting

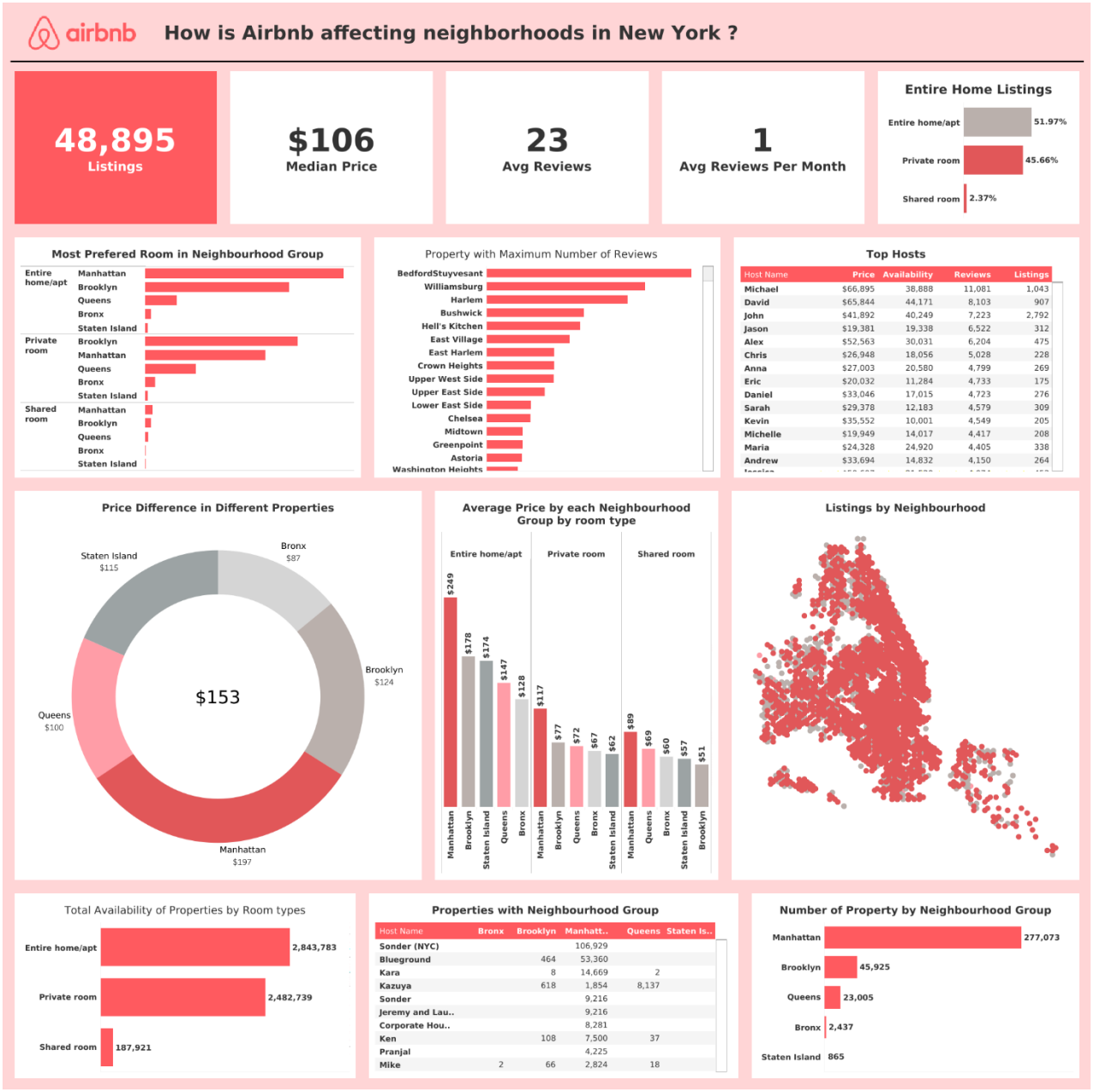
Reporting is a most important and underrated skill of a data analytics field. Because being a Data Analyst you should be good in the easy and self-explanatory report because your model will be used by many stakeholders who are not from a technical background. a) High-Level Design Document (HLD) b) Low-Level Design Document (LLD) c) Architecture d) Wireframe e) Detailed Project Report f) PowerPoint Presentation

5. Modelling Data

Modelling is the process of analyzing the data objects and their relationship to the other objects. It is used to analyze the data requirements that are required for the business processes. The data models are created for the data to be stored in a database. The Data Model's main focus is on what data is needed and how we have to organize data rather than what operations we have to perform.

6. Deployment

We created a Tableau Dashboard


airbnb
How is Airbnb affecting neighborhoods in New York ?
48,895
Listings
$106
Median Price
23
Avg Reviews
1
Avg Reviews Per Month
Entire Home Listings
Entire home/apt
51.97%
Private room
45.66%
Shared room
2.37%
Most Prefered Room in Neighbourhood Group
Entire home/apt
Manhattan
Brooklyn
Queens
Bronx
Staten Island
Private room
Brooklyn
Manhattan
Queens
Bronx
Staten Island
Shared room
Manhattan
Brooklyn
Queens
Bronx
Staten Island
Property with Maximum Number of Reviews
BedfordStuyvesant
Williamsburg
Harlem
Bushwick
Hell's Kitchen
East Village
East Harlem
Crown Heights
Upper West Side
Upper East Side
Lower East Side
Chelsea
Midtown
Greenpoint
Astoria
Washington Heights
Top Hosts
Host Name Price Availability Reviews Listings
Michael $66,895 38,888 11,081 1,043
David $65,844 44,171 8,103 907
John $41,892 40,249 7,223 2,792
Jason $19,381 19,338 6,522 312
Alex $52,563 30,031 6,204 475
Chris $26,948 18,056 5,028 228
Anna $27,003 20,580 4,799 269
Eric $20,032 11,284 4,733 175
Daniel $33,046 17,015 4,723 276
Sarah $29,378 12,183 4,579 309
Kevin $35,552 10,001 4,549 205
Michelle $19,949 14,017 4,417 208
Maria $24,328 24,920 4,405 338
Andrew $33,694 14,832 4,150 264
Price Difference in Different Properties
Staten Island
$115
Bronx
$87
Brooklyn
$124
Queens
$100
$153
Manhattan
$197
Average Price by each Neighbourhood Group by room type
Entire home/apt
Manhattan $249
Brooklyn $178
Staten Island $174
Queens $147
Bronx $128
Private room
Manhattan $117
Brooklyn $77
Queens $72
Bronx $67
Staten Island $62
Shared room
Manhattan $89
Queens $69
Bronx $60
Staten Island $57
Brooklyn $51
Listings by Neighbourhood
Total Availability of Properties by Room types
Entire home/apt
2,843,783
Private room
2,482,739
Shared room
187,921
Properties with Neighbourhood Group
Host Name Bronx Brooklyn Manhatt.. Queens Staten Is..
Sonder (NYC) 106,929
Blueground 464 53,360
Kara 8 14,669 2
Kazuya 618 1,854 8,137
Sonder 9,216
Jeremy and Lau.. 9,216
Corporate Hou.. 8,281
Ken 108 7,500 37
Pranjal 4,225
Mike 2 66 2,824 18
Number of Property by Neighbourhood Group
Manhattan 277,073
Brooklyn 45,925
Queens 23,005
Bronx 2,437
Staten Island 865