|  |
| --- |
| **Airbnb Market Analysis** |
| September 2023 – December 2023  BY: ANUSHKA PATIL |

**ABSTRACT**

This project delivers a comprehensive data-driven analysis of Airbnb’s impact on U.S. housing markets using advanced data analytics techniques. By leveraging Python, SQL, and Tableau, the platform evaluated over 2 million listings across major cities and uncovered strong correlations (up to 80%) between listing saturation and rental price volatility. The findings support municipal housing policy, investor strategies, and competitive benchmarking in real estate. The solution focuses on actionable insights and interactive visual storytelling to serve diverse stakeholders from public planning agencies to real estate firms.

**INTRODUCTION**

With the exponential rise of short-term rental platforms like Airbnb, understanding their influence on traditional housing supply and rental dynamics has become critical. This project investigates the underlying economic ripple effects caused by listing density and pricing variability. The system is built on a modern data analytics stack combining data engineering, statistical modeling, and business intelligence, designed for both exploratory and policy-oriented use.

**ROLE & TOOLS USED**

**Role:** Data Analyst, Dashboard Designer, Project Lead

**Focus Areas:** Housing Market Analytics, Data Engineering, Real Estate Economics, Stakeholder Reporting

• **Data Pipeline & Storage:** Built a robust ETL pipeline using Python (Pandas, NumPy) and PostgreSQL to ingest and transform Airbnb listing data, housing datasets, and economic indicators from open government APIs.

* **Exploratory Data Analysis (EDA):** Applied correlation matrices, regression models, and clustering to identify patterns between Airbnb penetration and local rental price shifts.
* **Visualization & Insights:** Designed interactive dashboards in Tableau to present trends by city, zip code, and listing type. Integrated filters for monthly price, availability, and host type to support dynamic exploration.
* **Advanced SQL & Queries:** Used PostgreSQL for querying millions of records efficiently, supporting join operations across spatial datasets (zoning, census, rent control).
* **Statistical Modeling & Forecasting:** Employed Python's stats models and scikit-learn to forecast the impact of listing growth on median rents using time-series forecasting and linear regression.
* **Project Management & Reporting:** Coordinated stakeholder updates using Agile methodology. Created biweekly reports and executive summary decks highlighting actionable insights for urban planners and investors.
* **Tech Stack & Tools:** Python, SQL (PostgreSQL), Tableau, Excel, Git, Google Collab, Jupiter Notebooks, Pandas, NumPy, scikit- learn, stats models.
* Designed monthly market impact forecast models to support policy and business decisions.

**RESULTS & ACHIEVEMENTS**

• Identified 80%+ correlation between Airbnb listing density and rental inflation in 5 major cities.

• Designed 3 interactive Tableau dashboards used by 10+ stakeholders.

• Enabled real-time filtering and drill-down by neighborhood, host type, and availability.

• Final report submitted to university research board for consideration in city-level planning studies.

**LEARNINGS & REFLECTION**

• Developed deep proficiency in end-to-end data analytics lifecycle.

• Learned stakeholder reporting through clear visualizations and agile documentation.

• Strengthened expertise in market analysis, predictive modeling, and geo-spatial analytics.

• Gained hands-on experience with scalable databases, BI dashboards, and Python-based EDA tools.

**PROBLEM STATEMENT**

* Lack of structured data linking Airbnb expansion to local housing economics.
* Absence of predictive models to estimate long-term rental market distortions.
* Difficulty for cities and investors to navigate short-term rental impact across neighborhoods.
* Need for real-time visualization tools tailored for decision-makers and analysts.

**APPROACH / SOLUTION**

* + Collected large-scale datasets from Airbnb, Zillow, and U.S. Census via APIs.
  + Built a centralized data warehouse using PostgreSQL and cleaned datasets using Python scripts.
  + Applied linear regression and k-means clustering to profile high- impact zones.
  + Created Tableau dashboards for rental trends, listing penetration, and host activity analysis.

**LINK OF THE PROJECT**

[**https://public.tableau.com/app/profile/anushka.patil6343/viz/AirbnbMarketAnalysisforNYC/Dashboard1**](https://public.tableau.com/app/profile/anushka.patil6343/viz/AirbnbMarketAnalysisforNYC/Dashboard1)

A screenshot of a computer screen

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated