Report

One Day Hotel Revenue & Booking Insights Dashboard

📌 Project Objective and Business Context

The project is designed to analyze a single day's worth of hotel booking activity (03 August 2016) across five major cities (Miami, New York, Los Angeles, San Francisco, Chicago), for five distinct hotel brands. Management can use this report for tactical decisions on pricing, marketing, and operations for peak days.

📊 Exploratory Data Analysis & Audit

The dataset contains 100 rows, each representing a single room booking. Key raw features reviewed:

- **HotelID/HotelName/Location**: Identification of unique hotels and relation to city for market segmentation.
- **RoomType**: Categorical feature with four values: Deluxe (most frequent), Suite, Double, Single.
- **Rating**: Discrete integer, range confirmed as 3-5, mean of 4.01.
- **PricePerNight**: Ranged from 50 to 250, mean at 160.53.
- **BookingStatus**: Four statuses: Confirmed, Cancelled, Pending, Not Defined.
- **CheckInDate**: All records for 03 August 2016, confirming a single-day slice.

🧹 Data Cleaning: Column-by-Column Actions

Hotel Name/Location/ Room Type:

- Ensured consistency of text (e.g., no "deluxe", "Deluxe" mismatch).
- Verified no trailing/leading spaces.
- Mapped all records to one of the five brands and five locations.

Rating:

- Confirmed all entries are valid integers in 3-5 range.
- No nulls or out-of-range points.

PricePerNight:

- Ensured all prices were positive and within typical hotel range.
- Handled missing prices (none in this dataset).
- Detected and reported any outliers (none found).

BookingStatus:

- 'Not Defined' status flagged for potential data or process issues.
- Converted all to capitalized form for easy grouping.

CheckInDate:

- Confirmed uniformity (all "03 August 2016").
- Validated for system compatibility by converting to ‘yyyy-mm-dd’ if needed for ETL processes.

Duplicates:

- Checked all fields for duplicates; none found.

🔧 Data Transformation: Analytical Engineering Steps

Aggregations:

- Grouped by RoomType, Location, and HotelName for:
  - Count of bookings
  - Sum and mean of PricePerNight
  - Sum of Ratings

New Features:

- Calculated `TotalRevenue` per hotel and per location.
- Derived `GuestCount` as 1 per booking, then summed for totals.
- Generated categorical encodings for BookingStatus for visual pipeline.

Pivot Tables:

- Created RoomType by Location and BookingStatus by Count matrices, supporting heatmaps and bar charts.
- Pivoted hotel-level sums for dashboard metrics.

📈 Dashboard Metrics Interpretation

Main KPIs:

- **Total Price:** 15,000 (indicates strong daily revenue).
- **Average Price/Rating:** 160.53 and 4.01—competitive positioning.
- **Total Guests:** 100, matching row count.

Booking Status:

- **Confirmed (39%)** — healthy, but improvement needed.
- **Cancelled (29%)** — signals risk in retention or customer service.

- **Pending (21%)** — outstanding potential or operational bottleneck.
- **Not Defined (11%)** — requires process audit.

RoomType Demand:

- Highest: Deluxe (40), then Suite, Double, Single, reflecting customer preference or availability.

Financial Performance by Location:

- Highest revenue: New York and Miami (approx 3.5K each), indicating successful geographic strategies.
- Lower-performing: San Francisco and Los Angeles (~2.8K).

Hotel-Level Analysis:

Hotel	Total Rating	Total Revenue
Mountain Inn	105	4.5K
Royal Stay	84	2.8K
Grand Palace	77	4.4K
City Lights Hotel	74	2.3K
Ocean View	61	2.3K

Mountain Inn and Grand Palace have the highest combined ratings and revenue.

✅ Recommendations & Business Actions

- **Increase Suite and Deluxe Inventory** in Miami and New York for peak periods.
- **Investigate High Cancellations**, especially for Double/Single rooms, to reduce revenue loss.
- **Automate Booking Status Definitions**, reducing 'Not Defined' instances and improving data quality.
- **Enhance Offers or Loyalty for Repeat Cities**, targeting cities with lower price/revenue.
- **Monitor Guest Ratings**: Focus on hotels/rooms with sub-4 average to boost long-term reputational value.

🔍 Additional Insights

- **Data Quality Issues** are minimal, but status/process automation improvements are recommended.
- **Pipeline Readiness**: Data, once cleaned and transformed as described, can be directly used for automated dashboards, real-time BI, or performance reviews.
- **Actionability**: Output allows for dynamic pricing, marketing campaigns for targeted locations, and operational process refinement.[1]