Hotel Booking Cancellation Analysis

**Summary**

This project aims to analyze hotel booking data to understand booking trends, cancellation behavior, customer preferences, and pricing patterns. Using Python (Pandas, Matplotlib, Seaborn), we explored data from city and resort hotels to extract insights that can help improve customer retention and business strategies.

**Problem Statement**

Many hotel bookings are canceled before check-in. These cancellations affect hotel revenue, resource planning, and customer satisfaction.  
**Key questions we tried to solve:**

* What percentage of bookings get canceled?
* How does ADR (Average Daily Rate) vary between canceled and not canceled bookings?
* Which countries are most involved in cancellations?
* What patterns are visible in booking trends, lead time, and stay duration?
* How does cancellation affect revenue across months?

**📊 Chart 1: Booking Cancellation Ratio**

**Chart:** (Insert the pie chart showing percentage of canceled vs not canceled bookings)

**Insight:**  
About **72% of the bookings were not canceled**, while **28% were canceled**. This shows a considerable cancellation rate that can significantly affect planning and profit margins.

**📊 Chart 2: Country-wise Booking Distribution**

**Chart:** (Insert the bar plot of top countries)

**Insight:**  
The majority of bookings come from **Portugal, the UK, and France**, but Portugal also shows a high number of cancellations, which is a red flag for the business.

**📊 Chart 3: Lead Time vs Cancellation**

**Chart:** (Insert the line or box plot showing lead time for canceled and non-canceled bookings)

**Insight:**  
**Canceled bookings tend to have much higher lead times**, indicating that people who book far in advance are more likely to cancel.

**📊 Chart 4: ADR per Month (for canceled bookings)**

**Chart:** (Insert barplot showing monthly ADR for canceled bookings)

**Insight:**  
ADR is generally higher for canceled bookings during peak months, suggesting that customers cancel when prices surge, possibly to rebook later at lower rates or elsewhere.

**📊 Chart 5: Average Daily Rate (ADR) Over Time**

**Chart:** (Insert line chart comparing ADR for canceled and non-canceled bookings)

**Insight:**  
**Non-canceled bookings maintain a stable ADR**, while **canceled bookings show fluctuations**, indicating revenue instability due to cancellations.

**📊 Chart 6: Distribution of Booking Channels**

**Chart:** (Insert pie or bar chart for booking channels like TA/TO, Direct, Corporate)

**Insight:**  
Most bookings come through **TA/TO (Travel Agencies/Tour Operators)**, which have higher cancellation rates compared to direct bookings.

**Conclusion**

* A significant portion of bookings gets canceled, especially those made far in advance.
* Portugal shows high cancellation rates, which may require targeted strategies.
* Higher ADR leads to more cancellations—pricing strategies should be reviewed.
* The data shows predictable trends that can help improve customer engagement, reduce cancellations, and optimize pricing.