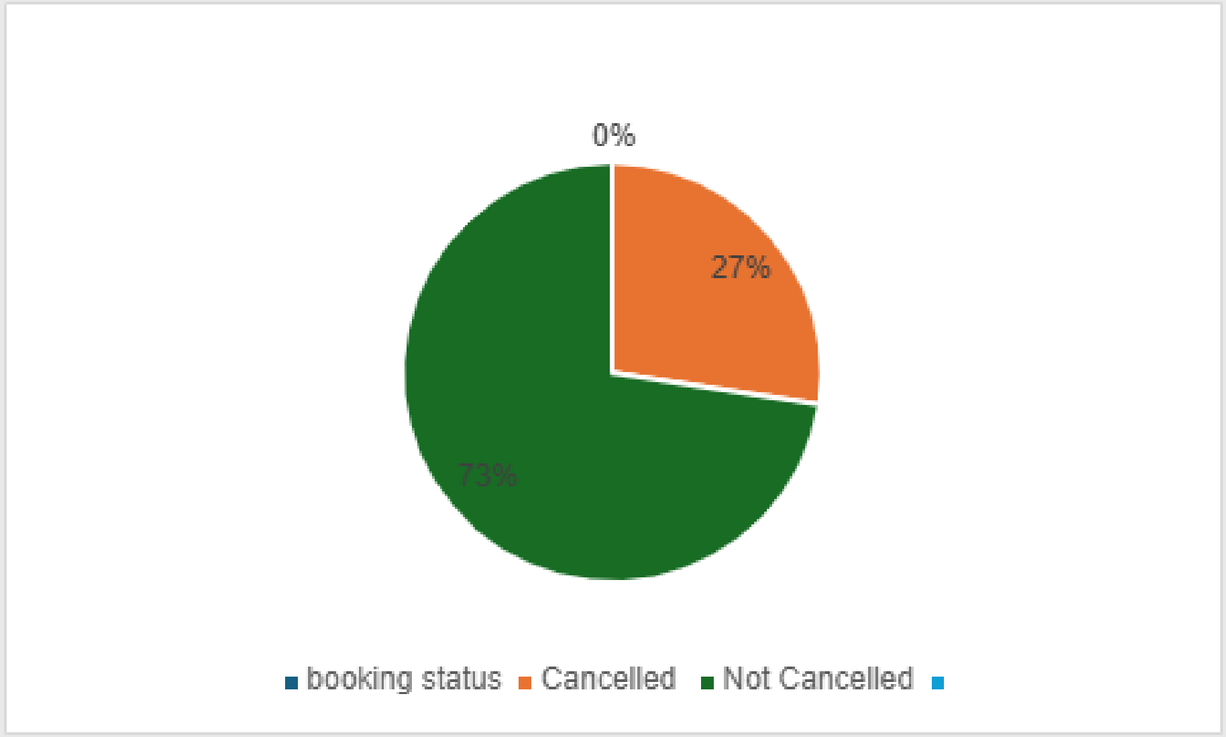
PRESENTATION

Hotel Booking Analysis

Booking Cancellation Rate



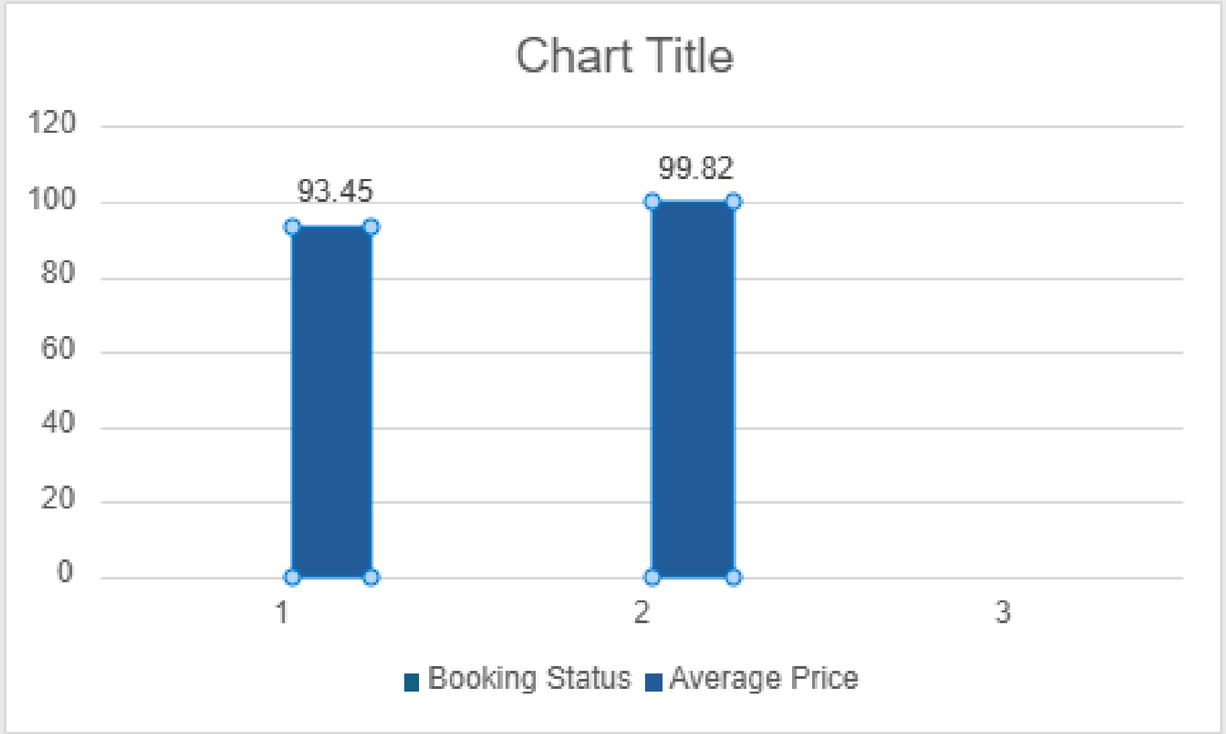
Canceled: 238 bookings

Not Canceled: 629 bookings

About 27% of bookings were canceled, which could indicate customer dissatisfaction, pricing issues, or .changes in travel plans

verage Room Price by Booking Status

Booking Status	Average Price
Canceled	9.345
Not Canceled	9.982



Canceled bookings had an average price of 93.45

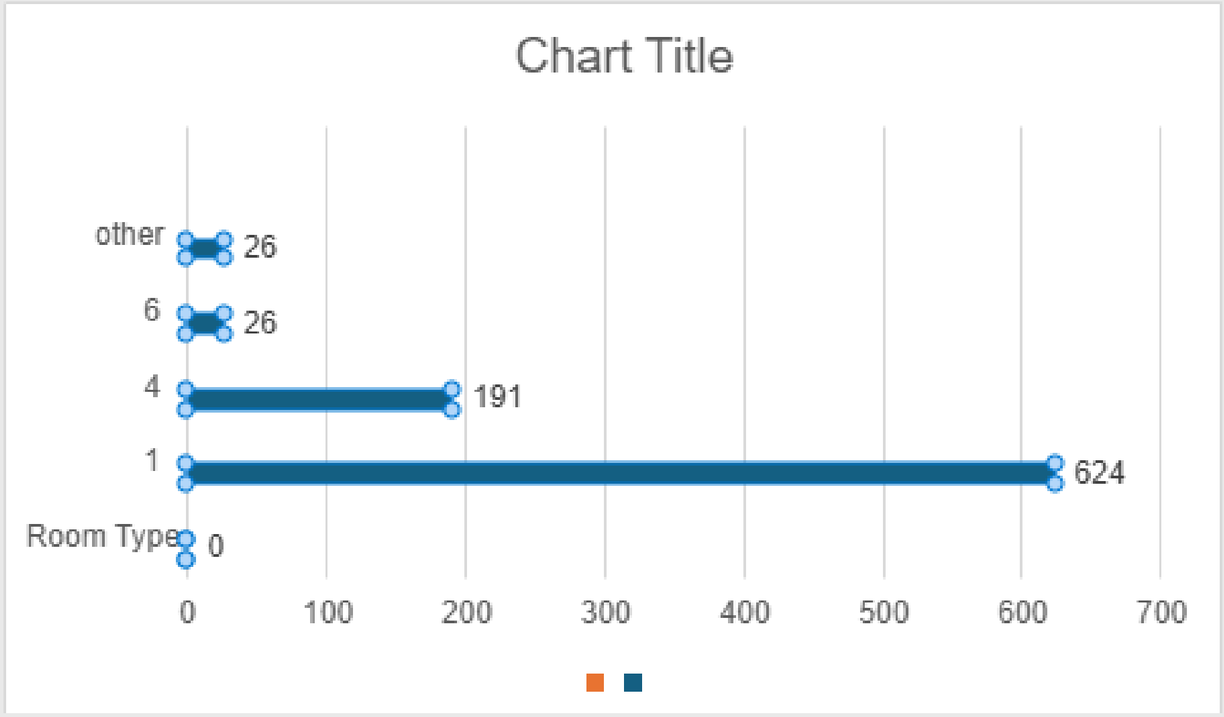
Not canceled bookings had an average price of 99.82

Guests who paid higher prices were less likely to cancel, suggesting that premium rooms may be more .valued or booked by more committed guests

Room Type Popularity

Room Type 1: 624 bookings

Room Type 4: 191 bookings



Room Type 6: 26 bookings

Other types: 26 bookings

Room Type 1 is by far the most popular, indicating it may offer the best combination of price, comfort, or availability.

Python Code

1. Importing Libraries

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

try:
```

Import essential libraries for data handling and visualization

Reading the Data

```
df = pd.read_csv('first_inten_project.csv')
```

Load the dataset from a CSV file

Renaming Columns

```
df = df.rename(columns={
    'booking status': 'Booking_Status',
    'average price': 'Average_Price',
    'room type': 'Room_Type',
    'lead time': 'Lead_Time'
})
```

Rename columns to clean and consistent names

Data Cleaning

```
df = df.dropna(subset=['Booking_Status', 'Average_Price', 'Room_Type', 'Lead_Time'])
df['Average_Price'] = pd.to_numeric(df['Average_Price'], errors='coerce')
df['Lead_Time'] = pd.to_numeric(df['Lead_Time'], errors='coerce')
```

Remove missing values and ensure numeric data types

Pie Chart

```
df['Booking_Status'].value_counts().plot.pie(...)
```

Shows the percentage of bookings that were canceled vs. not canceled

Bar Chart

```
df.groupby('Booking_Status')['Average_Price'].mean().plot(kind='bar', ...)
```

Compares the average room price for canceled and non-canceled bookings

Room Type Distribution

```
df['Room_Type'].value_counts().plot(kind='barh', ...)
```

Displays how many bookings exist for each room type

Lead Time Boxplot – Boxplot

```
sns.boxplot(x='Booking_Status', y='Lead_Time', ...)
```

Compares how early people book for canceled vs. non-canceled bookings