**Hotel Booking Analysis Project**

**Overview**

This project aims to process and analyze hotel booking data to identify trends, booking preferences, and cancellations. By leveraging big data technologies such as HDFS, Hive, and Spark, the analysis provides insights to optimize revenue and enhance guest satisfaction.

**Author**

Saron Yaya

**Initial work**

* [Portfolio Projects](https://github.com/Saron222/PortfolioProjects)

**Released on**

* GitHub

**My professional profile on LinkedIn**

* [My LinkedIn Profile](https://www.linkedin.com/in/saron-yaya/)

**Objective**

The primary objective of this project is to utilize big data tools to implement a data-focused solution for analyzing hotel booking data. The project aims to:

* Identify seasonal trends and booking patterns.
* Analyze customer segmentation and booking channels.
* Understand cancellation rates and reasons.
* Optimize revenue through dynamic pricing and revenue management strategies.

**Components Used**

* **HDFS**: Used for data storage and management of large datasets.
* **Hive**: Used for querying and analyzing the data stored in HDFS.

**Data Source**

The dataset consists of hotel booking records, including:

* Reservation details (arrival date, departure date, lead time).
* Guest information (customer type, market segment).
* Booking channels and cancellation status.
* Revenue-related data (average daily rate).

**Data Operations and Transformations**

1. **Data Cleaning**: Removal of null values, duplicates, and irrelevant data.
2. **Data Aggregation**: Grouping and summarizing data to identify key trends and patterns.
3. **Segmentation Analysis**: Classifying customers based on booking behavior and preferences.
4. **Cancellation Analysis**: Investigating cancellation trends and reasons.
5. **Revenue Management**: Applying dynamic pricing strategies and optimizing room inventory.

**Findings**

* **Seasonal Trends**: Identified peak and off-peak booking seasons.
* **Customer Segmentation**: Segmented customers into distinct groups based on booking behavior.
* **Cancellation Trends**: Analyzed the rates and reasons for cancellations.
* **Revenue Insights**: Provided insights for optimizing pricing strategies and revenue management.

**Implications**

The insights derived from this analysis can help hotels:

* Optimize pricing and resource allocation based on seasonal trends.
* Personalize marketing campaigns and services for different customer segments.
* Reduce cancellations and revenue loss by understanding cancellation trends.
* Enhance guest satisfaction and loyalty through better service.

**Future Directions**

* **Predictive Analytics**: Implement predictive models for booking demand and revenue forecasting.
* **Geospatial Analysis**: Analyze the geographic distribution of bookings to optimize marketing strategies.
* **Dynamic Pricing**: Develop and implement dynamic pricing algorithms for real-time rate adjustments.