



**M.KUMARASAMY**  
**COLLEGE OF ENGINEERING**

**NAAC Accredited Autonomous Institution**

Approved by AICTE & Affiliated to Anna University

ISO 9001:2015 Certified Institution

Thalavapalayam, Karur, Tamilnadu.



**ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

# PREDICTIVE MODELLING ANALYTICS

**18AIC305JT**

# HOTEL BOOKING CANCELLATION FORECAST

GUIDED BY,  
**SHUBA SRI S IBM - Corporate Trainer**

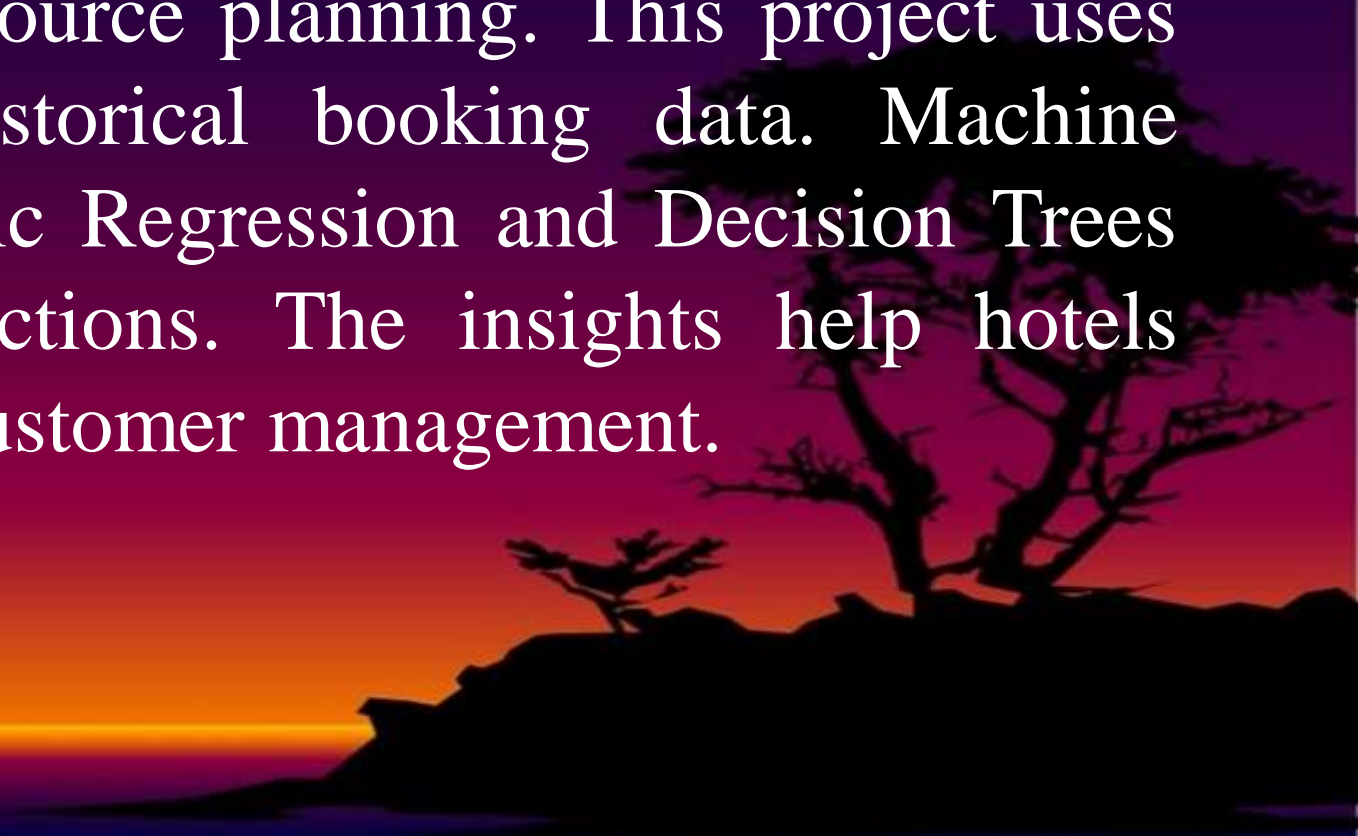
PRESENTED BY,  
NITHISHKUMAR S 927622BAL034  
SRIRAM R 927622BAL044  
YADHAVARAMANAN C 927622BAL055

# ABSTRACT

Hotel booking cancellations impact revenue and resource planning in the hospitality industry. This project uses IBM SPSS to develop a predictive model for forecasting cancellations based on historical data. Key factors such as booking lead time, customer type, and previous cancellations are analyzed. Machine learning techniques like Logistic Regression and Decision Trees are applied to improve accuracy. The model helps hotels optimize strategies, reduce losses, and enhance customer satisfaction.

# PROBLEM STATEMENT

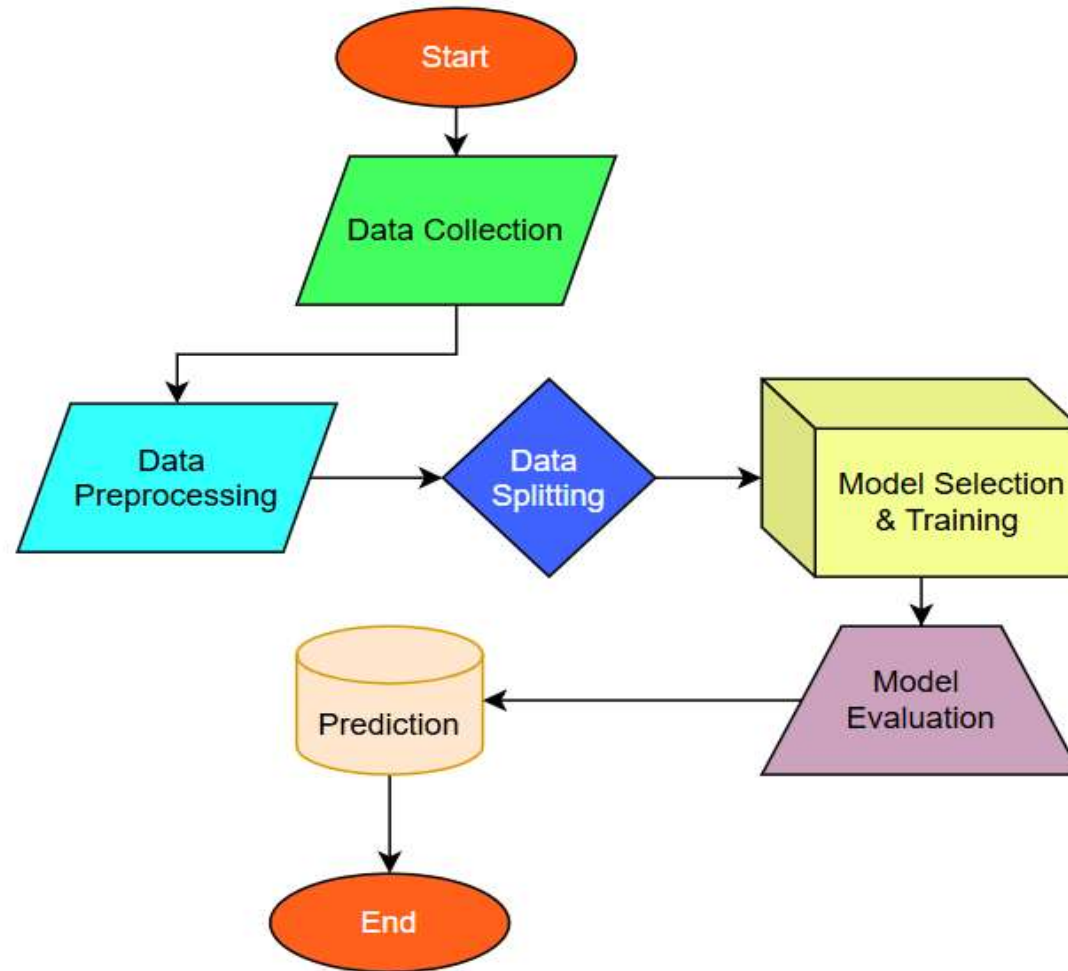
Hotel booking cancellations lead to revenue loss and operational inefficiencies. Predicting cancellations helps hotels optimize occupancy and resource planning. This project uses IBM SPSS to analyze historical booking data. Machine learning models like Logistic Regression and Decision Trees are applied to make predictions. The insights help hotels reduce losses and improve customer management.

A silhouette of a tree is visible on the right side of the slide, set against a background of a sunset or sunrise with a gradient from orange to purple.

# EXISTING SYSTEM



# PROPOSED SYSTEM





# TOOLS USED

- ❑ IBM SPSS
- ❑ Dataset
- ❑ Data Preprocessing
- ❑ Machine Learning Models
- ❑ Evaluation Metrics



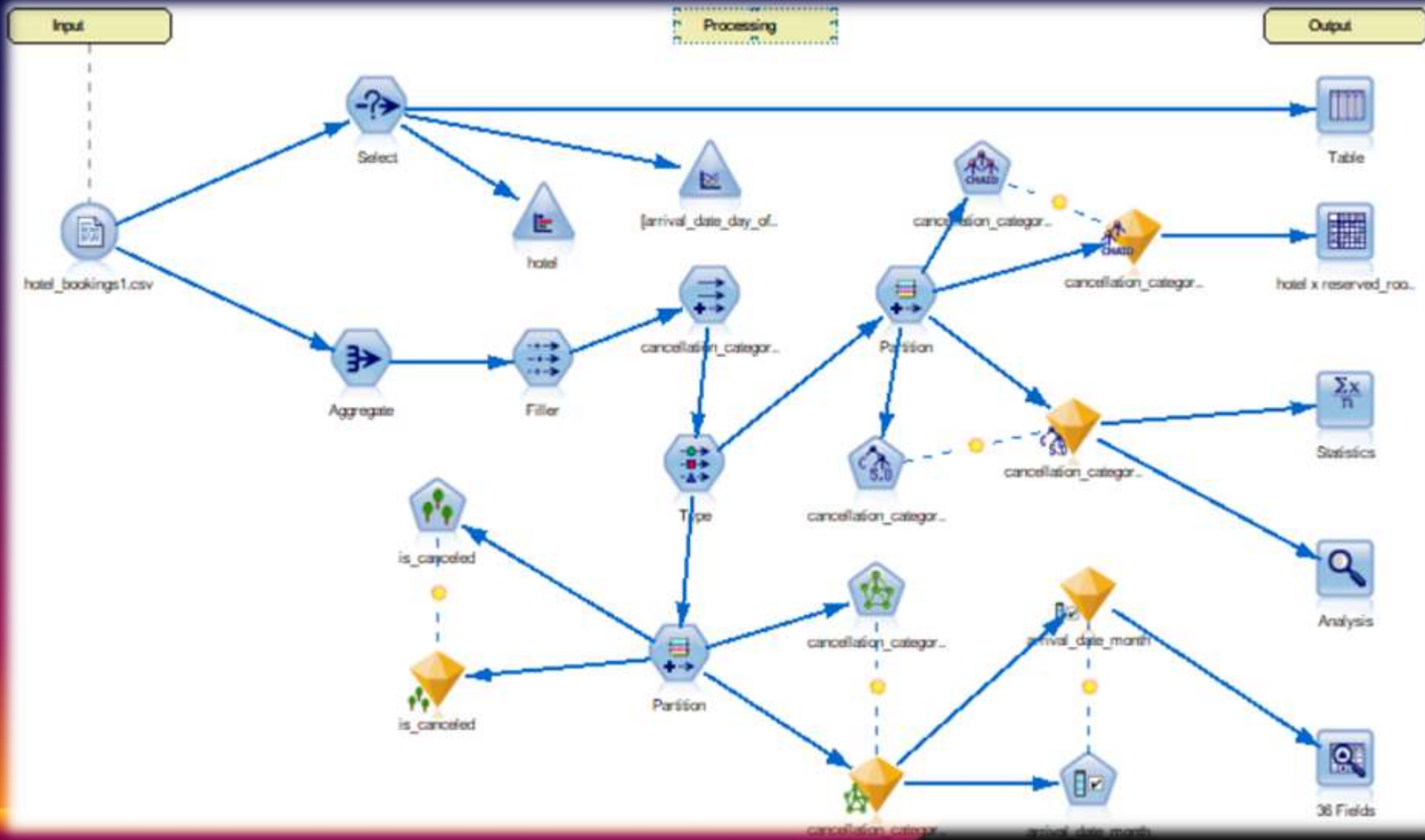
# MODEL DETAILS

- ❑ Logistic Regression
- ❑ Decision Tree
- ❑ Model Evaluation
- ❑ Cross-validation





# Stream



# Sample Prediction

## Classification for cancellation\_category

Overall Percent Correct = 100.0%

Observed	Predicted		Row Percent
	Canceled	Not Canceled	
Canceled	100.0%	0.0%	100.00
Not Canceled	0.0%	100.0%	80.00

Distribution of hotel

File Edit Generate View

Table Graph Annotations

Value	Proportion	%	Count
Royal Palace		19.13	463
Mountain Lodge		20.45	495
Grand Plaza		20.66	500
Ocean View Resort		19.88	481
Sunset Inn		19.88	481

OK

# Prediction

Matrix of hotel by reservation\_status

File Edit Generate

Matrix Appearance Annotations

reservation\_status

hotel	Canceled	Check-Out	No-Show
Grand Plaza	17332	30041	524
Mountain Lodge	17273	29968	460
Ocean View Resort	17095	29952	451
Royal Palace	17114	30292	474
Sunset Inn	17220	30079	505

Cells contain: cross-tabulation of fields (including missing values)

Chi-square = 10.498, df = 8, probability = 0.232

OK

Analysis of [cancellation\_category]

File Edit

Analysis Annotations

Collapse All Expand All

Results for output field cancellation\_category

Overall Results

Comparing \$C-cancellation\_category with cancellation\_category

'Partition'	1_Training		2_Testing	
Correct	119,693	100%	51,513	100%
Wrong	0	0%	0	0%
Total	119,693		51,513	

Output field cancellation\_category, splitting by field cancellation\_category

cancellation\_category = Canceled

Comparing \$C-cancellation\_category with cancellation\_category

'Partition'	1_Training		2_Testing	
Correct	34,457	100%	14,966	100%
Wrong	0	0%	0	0%
Total	34,457		14,966	

cancellation\_category = Not Canceled

Comparing \$C-cancellation\_category with cancellation\_category

'Partition'	1_Training		2_Testing	
Correct	85,236	100%	36,547	100%
Wrong	0	0%	0	0%
Total	85,236		36,547	

OK

# Prediction

## Random Trees

### Model Information

Target Field	is_canceled
Model Building Method	Random Trees Regression
Number of Predictors Input	1
Root Mean Squared Error	0.452
Relative Error	0.994
Variance Explained	0.006

The screenshot shows a software window titled "Statistics of [20 fields][is\_canceled]". It has a menu bar with "File", "Edit", and "Generate". Below the menu bar are tabs for "Statistics" and "Annotations". There are buttons for "Collapse All" and "Expand All". The main content area shows a tree view with the following data:

is_canceled	
Statistics	
Count	171206
Mean	0.289
Min	0
Max	1
Range	1
Variance	0.205
Standard Deviation	0.453
Standard Error of Mean	0.001

lead_time	
Statistics	
Count	171206
Mean	84.304
Min	0
Max	737
Range	737
Variance	8146.891
Standard Deviation	90.260
Standard Error of Mean	0.218

Pearson Correlations		
is_canceled	0.202	Strong

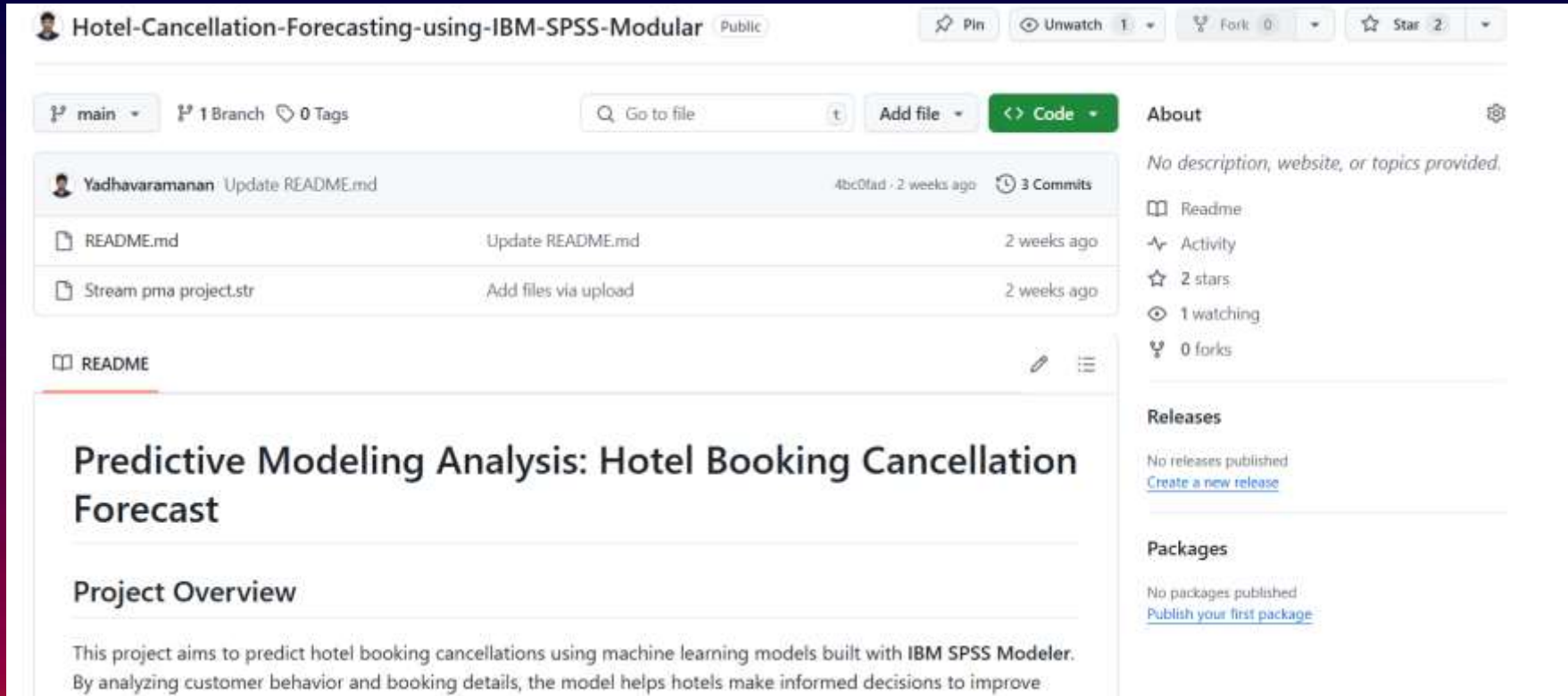
arrival_date_year	
Statistics	



# Conclusion

The project uses historical hotel booking data to predict cancellations, focusing on key factors such as **lead time**, **previous cancellations**, **customer type**, **deposit type**, and **market segment**. By analyzing these features, the predictive model helps forecast cancellations, allowing hotels to optimize their strategies, manage resources effectively, and reduce revenue loss. Machine learning techniques, including **Logistic Regression** and **Decision Trees**, are applied to improve prediction accuracy, ultimately enhancing customer satisfaction and operational efficiency.

# GitHub Link



The screenshot shows a GitHub repository page for a project titled "Hotel-Cancellation-Forecasting-using-IBM-SPSS-Modular". The repository is public and has 2 stars and 0 forks. The main branch is "main". The repository contains two files: "README.md" and "Stream pma project.str". The "README.md" file is the selected file, and its content is displayed below the file list. The content of the README.md file is as follows:

**Predictive Modeling Analysis: Hotel Booking Cancellation Forecast**

**Project Overview**

This project aims to predict hotel booking cancellations using machine learning models built with IBM SPSS Modeler. By analyzing customer behavior and booking details, the model helps hotels make informed decisions to improve

The right sidebar of the repository page shows the "About" section, which states "No description, website, or topics provided." It also shows the "Releases" section, which states "No releases published" and provides a link to "Create a new release". The "Packages" section also states "No packages published" and provides a link to "Publish your first package".

<https://github.com/Yadhavaramanan/Hotel-Cancellation-Forecasting-using-IBM-SPSS-Modular>



**THANK YOU**

A person wearing a dark suit is holding a silver smartphone. A bright blue, sparkling particle trail emanates from the screen of the phone, extending upwards and to the left. The background is out of focus, showing other people in a social setting with warm lighting.