

Predictive Modeling of Transformative Pilgrimage Patterns: A Demographic Analysis of Religious Tourists at the Basilica of Bom Jesus, Goa

Swizel Rosann Monteiro

Vinayak Naik

Jarret Stevan Fernandes

ABSTRACT

Religious tourism has emerged beyond pilgrimage and has morphed into spirituality, architectural significance, and cultural curiosity. Goa is an abode to multiple UNESCO World Heritage Sites which draws pilgrims and tourists alike, offering a unique blend of spiritual and cultural experiences that attracts both pilgrims and tourists. This study investigates the changing patterns of religious tourism at the Basilica of Bom Jesus during the 18th Exposition of the Sacred Relics of St. Francis Xavier.

Using a mixed-methods approach, the research analyzes visitor demographics, motivations, and participation trends through survey data, pilgrim footfall records, and digital engagement metrics. A key focus is placed on the role of online platforms in reshaping pilgrimage behavior and influencing travel decisions in the post-pandemic era.

To deepen insights and enable predictive understanding, the study employs supervised machine learning models to forecast event participation based on demographic and behavioral features. A multi-label classification framework was adopted to account for visitors attending multiple events. Among the models tested, Gradient Boosting achieved the highest predictive performance and was deployed locally via an interactive interface for potential real-world use.

The findings reveal key shifts in pilgrimage behavior, including a notable rise in digital pilgrimage engagement, multigenerational participation from diverse regions, and demand for budget-friendly accommodations. Additionally, the study reveals how large-scale religious events significantly stimulate the tourism economy. The research highlights the interplay between tradition and modernity, emphasizing the Basilica's dual appeal as a spiritual sanctuary and cultural landmark and demonstrates how data-driven modeling can support religious institutions and tourism authorities in anticipating visitor behavior, identifying trends, and planning inclusive, future-ready pilgrimage experiences.

Keywords: Religious Tourism, Pilgrimage, St. Francis Xavier Exposition, Machine Learning, Multi-label Classification

1. Introduction

1.1 Background

Religious tourism has taken place since the dawn of civilization. Pilgrims traveled to pay respect to their guardians throughout the world since time immemorial. Tourism and pilgrimage have merged since then. The importance of pilgrimage sites lies in the authenticity of religious practices, the architecture and their rituals, which makes these attractive to both religious as well as cultural tourists. The presence of tourists in sacred places promotes its essence of spreading beliefs and promoting cultural exchange (Tan, 2018). Communities welcome visitors as a way of sharing their faith and fostering global connections (Abiola-Oke & Osiobe, 2024).

Goa, a small state in India is famous not just for its beaches but also its numerous historic religious places like the UNESCO World Heritage site of Basilica of Bom Jesus, the Mahadeva Temple, and the Safa Mosque, which make religious tourism a significant part of its tourism sector. Tourists from across the globe visit these holy sites to grow spiritually and experience its culturally enriched heritage (Mukherjee et al.,2020)(Tan,2018). Known for its beautiful mixture of diverse faiths and traditions, all religions coexist harmoniously with one another. Beautiful white Christian churches stand tall near the elaborate temples of Hindu origin and gorgeously domed mosques of Muslims.

The Basilica of Bom Jesus that has the sacred relics of St. Francis Xavier attracts thousands of visitors daily. Approximately 7 million pilgrims had attended the 18th Exposition of the relics ([TOI, 2024](#)). This large crowd indicates the site's religious and cultural significance, making it an ideal place for analyzing the transformative pilgrimage patterns and religious tourism behavior. With the growing role of technology and data in understanding human movement and behavior, analyzing pilgrimage engagement patterns provides a foundation for sustainable tourism planning and spiritual engagement. This study focuses on understanding traditional perceptions of pilgrimage through the lens of data science, analyzing visitor profiles, motivations, and participation using supervised machine learning.

1.2 Relevant Work

Religious tourism, being one of the earliest forms of travel, has acted as a bridge between spirituality and cultural motives. While traditional pilgrimage, rooted in faith, continues to grow, modern visitors mostly seek more than just religious fulfillment (Marine-Roig, 2015). Many travelers now want to explore the cultural heritage embedded within these sacred sites (Wiśniewski, 2018). This confluence of religious and cultural tourism creates not only great opportunities but also thereby sets its own challenges for these destinations. Sacred places now engage diverse audiences, including non-believers and cultural tourists, by offering experiences

that go beyond religiosity (Collins-Kreiner, 2010). The catholic church has recognized these changing motivations, and has adapted its strategies to better engage visitors. By integrating art, architecture, music, and hospitality into its outreach efforts, the Church has prepared itself to connect with a broader audience and promote sacred sites as pastoral tourism products. The Church's artistic heritage plays a key role in this approach, offering travelers not only spiritual fulfillment but also cultural enrichment (Tan, 2018). Researchers suggest that Christian hospitality, coupled with well-preserved artistic and architectural traditions, fosters meaningful connections between tourists and religious communities (Abiola-Oke and Osioobe, 2024).

In recent years, digital platforms have been used for expanding the reach of religious tourism. Churches and pilgrimage sites now use websites, social media, and online campaigns to communicate faith and promote their destinations, bridging the gap that exists between spiritual and cultural engagement. Social media is now used as an aid for destination branding. Visitors share their experiences, through photos and videos on Instagram, Facebook, and other social networking sites and thus influence the way people choose their spiritual destinations (Agarwal et al., 2021). This shows the importance of applying social media strategies into tourism management to attract and engage potential visitors effectively.

The COVID-19 pandemic (Raj & Griffin, 2020) has brought several challenges to the religious tourism sector, by disrupting the traditional pilgrimage practices and also reshaping visitors' beliefs and behaviors. The government imposed restrictions on travel and mass gatherings, and hence many religious communities turned towards online platforms for these services. These led to new forms of spirituality and worship, creating alternative ways for religious engagement (Raj and Griffin, 2020). The pandemic also helped in understanding the various difficulties that were faced by travelers to religious sites, emphasizing the need for flexibility and innovation within the tourism industry to accommodate these changing expectations in visitors.

The shift towards digital participation has had several effects including travelers increasingly relying on social media for decision-making (Agarwal et al., 2021). The stories and experiences shared online shape the perceptions of visitors widely and play a pivotal role in influencing their travel choices, particularly in the spiritual tourism sector. Moreover, religious institutions and tourism managers must adapt to these strategies, balancing the traditional pilgrimage practices with new digital trends and the growing demand for cultural experiences.

This interconnected evolution of faith, culture, and digital engagement reflects the dynamic nature of modern religious tourism, requiring thoughtful management to meet the expectations of today's travelers.

1.3 Objectives

In this study an attempt is made to:

- Categorize pilgrims based on their demographics, visit history, and motivations.
- Analyze correlations between visitor types and their participation behaviors.
- Develop a supervised machine learning model that predicts event participation.



Fig.1.Pilgrims from all around the world visiting the Bom Jesus Basilica, Goa during the SFX Exposition in 2024

1.4 Research Questions and Hypothesis

- RQ 1. What are the primary factors influencing pilgrimage participation at the Basilica of Bom Jesus?
 - *H1. Demographic factors (age, gender, region) significantly influence the likelihood of participating in the Exposition.*
- RQ 2. What factors influence repeat pilgrimages to the Basilica?
 - *H2. First time visitors have different engagement motivations compared to long time devotees and frequent pilgrims.*
- RQ 3. How has COVID-19 pandemic transformed pilgrimage patterns at the 18th Exposition held at the Basilica?
- RQ 4. What role do YouTube and other platforms play in shaping travel decisions and pilgrim behavior?
- RQ 5. Can the surge in Goa's tourist arrivals in 2024 be statistically linked to the 18th Exposition?
 - *H3. The Exposition had no significant effect on the total number of tourist arrivals in 2024 compared to previous years*
- RQ 6. Can machine learning models effectively predict visitor event participation based on survey data?

By examining the above questions through survey data and observations, the study aims to understand how the motivations vary across the demographic groups. It seeks to uncover what are the reasons that drive pilgrims to return to the Basilica over the years. It also explores the shift in participation brought about by pandemic restrictions and evolving visitor preferences. With the growing importance of digital engagement, the study also investigates how online content influences visitor perceptions and participation in religious tourism (Dein and Watts, 2023). Furthermore, it examines whether the observed increase in tourist numbers during 2024 is associated with the 18th Exposition.

On the whole, through these questions, the research studies the interplay between traditional religious practices, modern trends such as cultural tourism, and the influence of digital platforms, offering insights to stakeholders for enhancing the visitor experience and ensuring sustainable tourism practices.

1.5 Scope

This study focuses on religious tourism in Goa, by specifically analyzing visitor engagement during the novena days for the 18th Exposition of the Sacred Relics of St. Francis Xavier held at the Basilica of Bom Jesus. The geographic scope is limited to the Basilica and its surrounding region, including online platforms that streamed the event. The research considers data collected during the 2024 Exposition, incorporating both in-person and online participants.

The study primarily investigates visitor demographics, motivations, frequency of visits, and the impact of social media platforms especially YouTube on travel decisions and engagement with the event. The research is limited to responses from approximately 2,000 survey participants, along with historical footfall statistics and digital engagement metrics. Using a mixed-methods approach, the study integrates quantitative data analysis and machine learning models to identify evolving pilgrimage patterns.

While the findings offer insights into the post-pandemic shifts in religious tourism, the study does not include financial or institutional aspects of event management, nor does it explore theological interpretations of pilgrimage. The scope is centered on behavioral, technological, and cultural dimensions of contemporary religious tourism in Goa.



Fig.2. Pilgrims attending Mass and Veneration of Relics at the Basilica

2. Methodology

2.1 Research Design

This study employs a quantitative, data-driven approach to understand and predict participation in events using survey data, attendance records, and YouTube viewership statistics. Both exploratory and predictive techniques were applied.

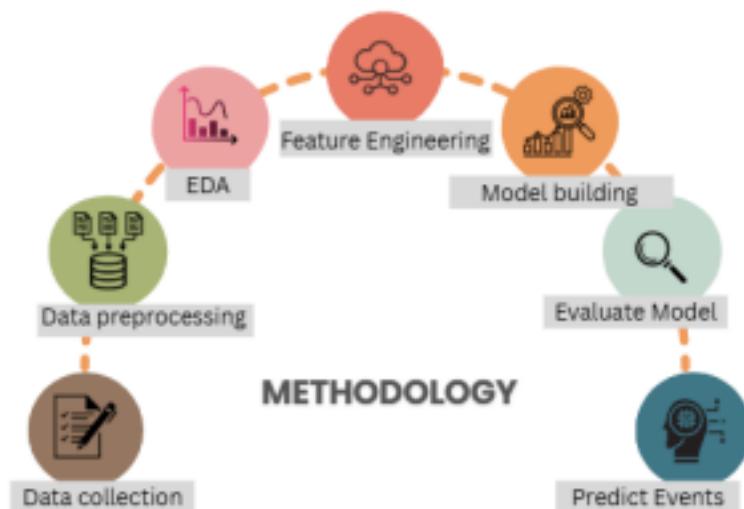


Fig.3. Flowchart of Methodological Approach

2.2 Data Collection

The research was conducted during the 18th Exposition of the Sacred Relics of St. Francis Xavier held starting from the inaugural day 21st of November, leading to the feast day on the 3rd of December. The research employs a mixed-methods approach, integrating both quantitative and qualitative methods to provide a comprehensive understanding of pilgrimage patterns, visitor demographics, and motivations at the Basilica of Bom Jesus during the Novena Days and the Exposition of the Sacred Relics of St. Francis Xavier. The data collection process involved survey questionnaires (Stamenkovic et.al, 2024), digital engagement metrics, and visitor footfall count to provide a holistic understanding of the visitors and their demographics at the event.

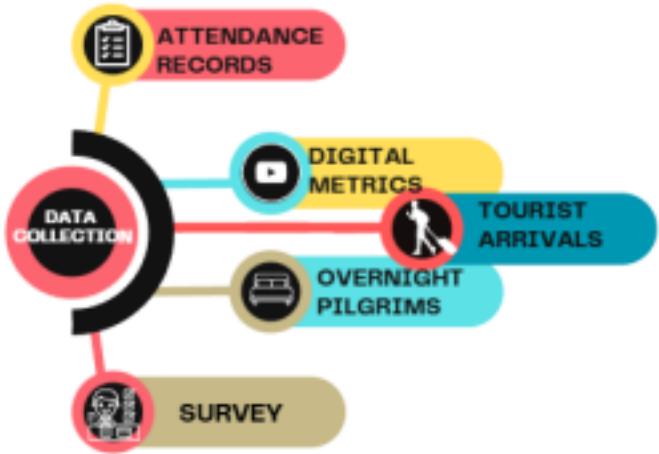


Fig.4. Sources of data collection

Attendance Records

To capture the scale and flow of participation during the event, data collection efforts focused on daily counts of attendees across various key points. This provided a detailed understanding of the turnout and logistics.



Fig.5 : Pilgrims attending mass services at the Basilica

Digital Metrics Collection

In addition to attendees, data on live-streamed Masses during the Novena days was collected from the YouTube channel ‘St Francis Xavier Official Basilica Bom Jesus Goa.’ The number of views for each Mass was recorded to assess digital engagement and participation. This data provided insights into how virtual religious services influence pilgrimage behavior, particularly

in the post-pandemic era.



Fig.6. YouTube Mass Services offered during the Exposition period

Overnight Pilgrims: Daily Check-in Analysis

We explored the daily check-ins of overnight pilgrims, categorized as Pilgrim Village guests and camping guests, during the exposition. The data provides insights into the accommodation preferences and their trends over the exposition period. Pilgrim Village Guests were individuals staying in designated accommodations provided near the Basilica and Camping Guests were pilgrims who opted for camping facilities available during the exposition.

The arrivals of walking pilgrims, who journeyed to the venue from various regions in India, were documented over the days leading up to the Feast Day. These pilgrims form an important part of the event's attendance, and their engagement was tracked to understand their participation in the religious activities and rituals.

The analysis showcases the influx of overnight pilgrims and highlights peak days for accommodations. This can be correlated with event schedules, such as major masses or other significant occasions or weekends.

Survey Design and Data Collection

To gain insights into the pilgrims' demographics, their motivations, a structured survey (Pillai et al., 2017) was designed and conducted during the Novena Days at the Exposition of the Sacred Relics of St. Francis Xavier at the Basilica of Bom Jesus. The survey included questions related to visitor demographics, event engagement, pilgrimage history, and experiences, with a focus on the walking pilgrims too.

The survey was conducted in person by visiting the Basilica during Novena days. Pilgrims attending the services or visiting the site were approached and asked to complete the survey. These were then manually added to the database. Additionally, the Pilgrim Village which is a temporary settlement built by the government to accommodate pilgrims attending the Exposition was visited. This included barracks and tents that were set up for visitors and walking pilgrims respectively. Pilgrims also shared their experiences, especially those who undertook walking pilgrimages, providing valuable insights into traditional practices and its challenges. To further expand participation and maximize responses, the survey was also distributed online via a QR code link to a Google Form.



Fig.7: Pilgrims answering the survey at the Basilica

These online forms enabled outreach to individuals who we may not have been able to reach physically at the Basilica during the Exposition. A total of 2000 survey responses were collected from pilgrims attending the Exposition, representing a stratified sample from the estimated million attendees.



Fig.8. Poster featuring QR Code for Online Survey Participation

2.3 Data Preprocessing

Approximately 2000 responses were collected from participants through a structured questionnaire after which several preprocessing steps were undertaken to ensure the dataset's consistency, accuracy, and suitability for analysis.

Data Cleaning

The data was cleaned to ensure the reliability and accuracy of the findings. Attendance records and digital metrics from YouTube live-streamed Masses were reviewed and cleaned, ensuring precise measures of virtual engagement. Accommodation data, including daily check-ins at the Pilgrim Village and camping facilities, was normalized to account for inconsistencies in reporting formats.

Survey responses collected via Google Forms were exported into CSV format and cleaned using Microsoft Excel tools. This involved identifying and removing duplicate entries, addressing incomplete responses, and standardizing text-based answers for consistency.

In addition to Excel-based processing, the dataset underwent further cleaning and transformation using Python. The following key steps were executed programmatically:

Handling Missing Values

Missing or null values were handled through imputation and removal, ensuring a clean dataset for analysis.

Standardization

Standardize categorical text columns (e.g., 'State/Country', 'Gender') by converting to lowercase and stripping extra spaces.

Creation of new features

To facilitate deeper insights into visitor demographics and behavior, two new features were created:

- Age Group Classification: The existing Age data was categorized into distinct life stages to understand patterns across age segments.
- Visitor Frequency/ Type of Visitor: The 'How many years (approx.) have you been

'visiting the Basilica?' feature was used to identify the nature of the visitor's engagement with the destination.

These engineered features provide a more nuanced understanding of both the demographic and behavioral aspects of tourist visitation.

Encoding Categorical Variables

Categorical variables such as 'Gender:', 'State/Country', 'Are you camping at the Basilica?', 'Age groups', 'Type of visitors', and 'walking pilgrim' were encoded into numeric formats using Label Encoder from `sklearn.preprocessing`.

2.4 Exploratory Data Analysis (EDA)

Data exploration was done using descriptive statistics, frequency distributions, and creating visualizations including bar charts, line graphs, and histograms to identify trends in attendance and viewership.

Excel and Power BI were used to get meaningful insights. In Excel, functions such as conditional formatting facilitated the organization and preliminary exploration of the data. Using Data Analysis Expression (DAX) queries, cumulative attendance trends, demographic segmentation, accommodation patterns, and digital engagement metrics were analyzed in depth. Interactive dashboards were created that provided visual representations of time-series data, attendance peaks, and correlations between pilgrims' arrival dates and event schedules.

The combination of Excel, Power BI, and Python ensured a robust analysis, enabling actionable insights into visitor patterns, engagement trends, and accommodation preferences.

2.5 Target Variable

The target variable represented the set of events attended by each pilgrim, such as Veneration of Relics, Mass, and Fair/Cultural Event/Sightseeing. Since a single pilgrim could participate in multiple events, the problem was structured as a multi-label classification task. To prepare the data accordingly, the original event-attendance column which contained comma-separated values was transformed using multi-label binarization. This process involved creating separate binary columns for each event, where a value of 1 indicated that the pilgrim attended the event and 0 indicated non-attendance. This approach preserved the multi-event nature of participation and enabled us to apply machine learning models that support multi-label outputs.

2.6 Feature Associations

To examine the factors influencing pilgrims' participation in different events at the Basilica of Bom Jesus, we analyzed the associations between event attendance and various categorical predictor variables, including socio-demographic and behavioral factors.

Correlation Analysis (Numerical Variables)

The initial step involved analyzing associations between continuous or ordinal variables, such as Age and Number of Visits. To assess the strength and direction of relationships among numerical variables Pearson correlation was used for normally distributed continuous variables, while Spearman correlation was considered for ordinal or non-normally distributed data. A correlation heatmap was generated to visualize associations.

Association Analysis (Categorical Variables)

Since the key predictors and the target (event attended) were categorical in nature, standard correlation measures like Pearson correlation were not appropriate. This is because Pearson assumes a linear relationship between continuous variables and treats numerical labels of categorical data as meaningful, which they are not. Label-encoded categories do not preserve ordinal or relational meaning and using them directly could result in misleading or spurious correlations.

To accurately measure associations between categorical features and event attendance, the following approach was adopted:

- Chi-Square Test of Independence: Applied to evaluate whether a statistically significant relationship exists between each predictor and the event(s) attended.
- Cramér's V: Used to quantify the effect size of each relationship. This normalized statistic ranges from 0 (no association) to 1 (perfect association), offering an interpretable measure of how strongly a predictor variable is associated with attendance at each event.

Visualization of Associations

A heatmap of Cramér's V values was generated to visualize the strength of association between each predictor variable and the individual events:

- Y-axis: Event types (Mass, Veneration of Relics, Fair/Cultural Event, etc.)
- X-axis: Categorical predictor variables (e.g., Age Group, Gender, Visitor Type, Camping Status)
- Color Intensity: Represents the strength of the association based on Cramér's V

This visual summary facilitated the identification of which features were most influential in predicting event participation.

Hypothesis Testing

To statistically validate assumptions from EDA and correlation analysis, hypothesis tests were conducted. These included chi-square tests for the associations between categorical variables.

2.7 Feature Engineering

Feature selection is a critical step in building an effective classification model. Relevant features were selected using a combination of domain knowledge, understanding of pilgrim behavior and religious tourism and from the statistical association analysis.

Selected Features

The following features were identified as significant predictors of event participation:

- Age group
- Gender
- State/Country of origin
- Type of visitor (e.g., First-time, Occasional, Frequent, Long-Time Devotee)
- Camping status at the Basilica
- Weekend indicator
- Feast Day indicator

Rationale for Selection:

Chi-Square Tests and Cramér's V were used to assess the strength of association between each feature and individual event categories. (For a complete overview of the association strengths between visitor characteristics and event participation, see Appendix III)

Features with higher Cramér's V values (typically above 0.1) were retained as they indicated moderate to strong associations. State/Country and Type of Visitor had notably high association scores with multiple events (e.g., Mass, Veneration of Relics). Domain expertise supported including Age and Camping status, as these characteristics are often linked with pilgrimage intensity and spiritual engagement. This structured approach ensured that only the most relevant and impactful features were included, improving model accuracy and interpretability.

2.8 Model Selection

To handle the multi-output classification problem where multiple labels of event participation needed to be predicted simultaneously several classification algorithms were selected. These models were chosen based on their compatibility with multi-label or multi-output tasks and their underlying learning mechanisms, which allow for flexibility in capturing relationships across diverse features and target variables. Fig.9. illustrates the various classification algorithms evaluated during the model selection process to determine the most suitable one for event

prediction.



Fig.9. Classification Models considered for Selection

Random Forest Classifier

The Random Forest Classifier is an ensemble method that builds multiple decision trees from bootstrapped samples of the training data. By aggregating the predictions of individual trees, the model effectively reduces variance and enhances its ability to capture complex, non-linear relationships across the feature space. In this study, the model was wrapped using Scikit-learns Multioutput Classifier, enabling it to concurrently predict multiple target variables. Random Forest is advantageous due to its inherent feature selection capability and resistance to overfitting, making it a natural choice for multi-dimensional prediction tasks.

Decision Tree Classifier (with One-vs-Rest strategy)

The Decision Tree Classifier builds a tree-like structure where each node represents a feature split chosen based on criteria such as information gain or Gini impurity. In the multi-output scenario, the classifier was implemented within a One-vs-Rest Classifier wrapper nested in Multioutput Classifier. This configuration trains individual trees for each target variable independently. While decision trees provide clear interpretability through their hierarchical splitting, they can be prone to overfitting if not properly pruned or regularized, an important consideration for datasets with diverse label distributions.

Logistic Regression (with One-vs-Rest strategy)

Logistic Regression models the probability of class membership for each output by employing a logistic function. For multi-output classification, a One-vs-Rest Classifier was used to build separate binary classifiers for each target variable, encapsulated by the Multi-Output Classifier. This approach maintains model simplicity and interpretability, especially when the decision boundaries are expected to be linear. However, Logistic Regression faces challenges in capturing intricate, non-linear dependencies among features, which needs to be considered when selecting models for complex datasets.

Support Vector Machine (SVM) Classifier

Support Vector Machines aims to find the optimal separating hyperplane that maximizes the margin between the classes. In multi-label scenarios, the One-vs-Rest approach is utilized to extend SVM capabilities across all the target outputs by training one classifier per class. Within the Multi Output Classifier framework, in this arrangement each SVM independently learns a decision boundary tailored to a specific label. Despite its theoretical robustness, the application of SVMs in a multi-output setting can be computationally intensive, and careful parameter tuning is often required to achieve balanced classification.

k-Nearest Neighbors (k-NN) Classifier

The k-NN algorithm takes the majority vote of the k closest examples in the feature space to make predictions. Implemented using the Multi Output Classifier, the k-NN model directly applies its proximity-based decision rule to each of the target variables. Since its non-parametric nature no explicit training phase is involved, which is advantageous because of its simplicity. However, this method is sensitive to the choice of k value, feature scaling, and the presence of irrelevant features; these factors need to be systematically addressed during preprocessing to ensure accurate multi-output predictions.

Gradient Boosting Classifier

The Gradient Boosting Classifier is an ensemble technique that sequentially combines weak learners, typically decision trees, to minimize a differentiable loss function. This model iteratively focuses on the errors of the previous learners, refining its predictions with each successive tree. When used in a multi-output framework via the Multioutput Classifier, each target's prediction is improved iteratively by addressing the residuals from prior iterations. Gradient Boosting is valued for its flexibility and ability to fine-tune complex predictive relationships, though it requires careful management of hyperparameters and may involve increased computational demands.

2.9 Model Setup and Training

The classifier is set up to handle multiple target variables (e.g., predicting events for different categories). The dataset was preprocessed and split into features (X) and multi-label targets (y). A 70/30 train-test split was applied using `train_test_split()` from `sklearn.model_selection` to ensure that the models could be trained on 70% of the data and tested on the remaining 30%. This split provides a good balance between training performance and generalization capability on unseen data. All models were trained using multi-label classification techniques, primarily using Multi Output Classifier or One-Vs-Rest Classifier wrappers, as many standard classifiers are not inherently multi-label.

2.10 Model Evaluation

Each model's performance was assessed using multiple evaluation metrics, suitable for multi-label classification:

- Accuracy Score – Measures the proportion of correct predictions across all output labels.
- Precision, Recall, F1-Score – Computed for each class (event) using classification report to understand the model's per-label performance.
- Hamming Loss – Evaluates the average number of labels incorrectly predicted per instance. Lower Hamming Loss is better.
- Jaccard Score (macro) – Captures label-level similarity between predicted and true values, averaged across all labels. Higher means we have a better match.
- F1-Score (Macro)- Balance between precision & recall, averaged across labels.
- Samples avg F1-How well each individual pilgrim's labels were predicted.

These metrics allowed a comprehensive evaluation across all output dimensions and helped identify the most effective models for multi-event prediction.

F1 Score for Multi-label Classification

While Accuracy measures the overall percentage of correct predictions, it fails in imbalanced datasets or multi-label classification. A model might predict the majority class correctly, leading to high accuracy but poor performance on minority classes.

- Precision: Out of all predicted positive labels, how many were actually correct?
- Recall: Out of all actual positive labels, how many were correctly predicted?

Therefore, F1-score gives a better balance between false positives and false negatives, especially for multi-label event prediction, where a pilgrim can attend multiple events.

2.11 Model Testing

Following training and validation, all classifiers were tested on the hold-out test set (X_{test}) to evaluate generalization performance. The predicted labels (y_{pred}) were compared against the true labels (y_{test}) using the metrics mentioned above. This stage helped validate model robustness in real-world scenarios where first-time and recurring visitors participate in multiple religious and cultural activities.

2.12 Predictions

After training and validating the classification models, the final step involved generating predictions for unseen visitor data. A sample input was created with encoded demographic and visit-related features such as age group, gender, state or country of origin, type of visitor, and contextual details like whether the visit occurred during the weekend or on the Feast Day.

The model produced a multi-label output indicating the likelihood of participation in various event categories such as Fair/Cultural Event/Sightseeing, Veneration of Relics, and Mass. These predictions were decoded back into their original event labels for interpretability.

A utility function was also developed to handle raw input from new visitors, encode the inputs using Label Encoder, and decode the predictions from binary arrays to human-readable event names. This allows practical, real-time use of the model for forecasting visitor preferences and participation.

2.13 Deployment

To ensure the trained model can be reused and deployed for real-world applications, the final and best-performing classifier was serialized and saved using the joblib library. This facilitates portability and efficient storage of trained models.

- Model Serialization: The final Gradient Boosting model was saved using joblib to allow reuse without retraining.
- Prediction Interface: A command-line interface was created to interactively collect new visitor data and provide predictions. The interface includes input validation and real-time decoding of event participation results. To predict the events a visitor might attend, a custom prediction function was developed. It accepts new input data, encodes the categorical variables using pre-trained label encoders, and returns the predicted event(s) from the multi-output classifier.
- An interactive function was implemented to allow dynamic predictions by prompting the user for input values. This tool serves as a prototype for an intelligent assistant that could be used at the Basilica for visitor engagement. It takes human-readable input values (like "Yes", "No", or "First-Time Visitor") and encodes values using the same label encoders used during model training. It then outputs user-friendly event predictions
- Encoders and Metadata: All label encoders used during preprocessing were saved alongside the model to ensure consistent encoding for new predictions.
- Scalability Considerations: Although currently operating in a local environment, the model is structured to support future integration into a web or mobile application with real-time inputs.
- User Experience: A user-friendly interaction loop was designed to allow easy input from non-technical users, making the tool accessible to event organizers or help desk volunteers.

2.14 Ethical Considerations

The survey was conducted with informed consent from all participants, ensuring anonymity and confidentiality. Pilgrims were briefed about the purpose of the research, and participation was voluntary. Ethical clearance was obtained for the study to ensure compliance with research standards.

2.15 Software Tools and Libraries

Data analysis was conducted using Power BI and Python (Pandas, Scikit-learn, Seaborn, Statsmodels), with visualizations created using Matplotlib. Survey data was collected via Google Forms and exported as CSV.

Table 3.1 Software Tools and Libraries used

Category	Tool / Library	Purpose
Programming Language	Python (3.x)	Primary language for ML and data processing
Development Environment	Visual Studio code / Collab	Interactive coding and experimentation
Data Handling	pandas	Data manipulation and preprocessing
	numpy	Numerical computations and array handling
Preprocessing	LabelEncoder, StandardScaler	Encoding categorical variables and scaling features
Model Training	scikit-learn	Machine learning models and utilities

	MultiOutputClassifier	Handling multi-label classification tasks
	OneVsRestClassifier	Binary relevance strategy for multi-label tasks
	RandomForestClassifier	Ensemble model – baseline
	DecisionTreeClassifier	Simple interpretable tree-based model
	LogisticRegression	Linear classifier – good for binary/multi-label classification
	SVC	Support Vector Machine model
	KNeighborsClassifier	Proximity-based learner
	GradientBoostingClassifier	Ensemble method – best performing in your case
Evaluation Metrics	classification_report	Precision, recall, F1-score
	accuracy_score	Overall correctness of predictions
	hamming_loss	Fraction of incorrect labels
	jaccard_score	Similarity measure for predicted vs actual labels
	f1_score	Harmonic mean of precision and recall
	roc_curve, auc	Plotting ROC curves and computing Area Under Curve

Visualization	matplotlib.pyplot	Creating plots such as ROC curves
	seaborn (optional)	Enhanced visualizations (e.g., heatmaps, style)
	Power BI	Interactive data dashboards and visualization for insights and sharing
Interface/ Deployment	Streamlit	Building an interactive web-based prediction interface
Model Serialization	joblib	Saving/loading trained models and preprocessing objects (encoders etc.)
Background Styling	base64	Encoding image for Streamlit app background

3. Results and Analysis

3.1 Pilgrim Footfall Analysis

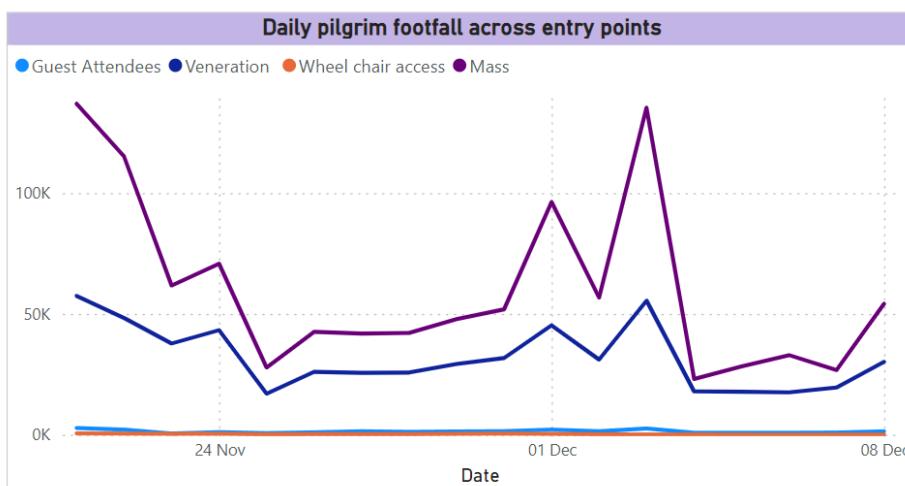


Fig.10. Daily pilgrim footfall across various entry points

The Basilica entrance had a higher footfall compared to the other entrances marked to be the primary gateway, highlighting its central role in visitor access to masses. There was a surge in attendance during the inaugural ceremony of the Exposition of the Sacred Relics of St. Francis Xavier held on the 21st of November where the relics were carried in a procession to the Sé Cathedral in an electric carriage. After which a peak was observed on November 24th, marking the start to the Novenas. A decline in daily counts was noted thereafter, reflecting shifts in visitation preferences. Attendance spiked again on December 1st, likely driven by the weekend, and reached its highest on the Feast Day, December 3rd, when pilgrims flocked to the Basilica to pay homage to the saint. The wheelchair-accessible entrance recorded steady but limited use, indicating consistent demand for accessibility services. The Veneration entrance also showed attendance patterns very similar to the Basilica entrance, with some variations corresponding to the Novena masses. Apart from this, there were 1,200 walking pilgrims who journeyed from the neighboring states of Maharashtra and Karnataka, emphasizing their deep reverence for the saint. Temporary accommodations, including camps and the pilgrim village, were used extensively to cater to the influx of visitors.

3.2 YouTube Viewer Count

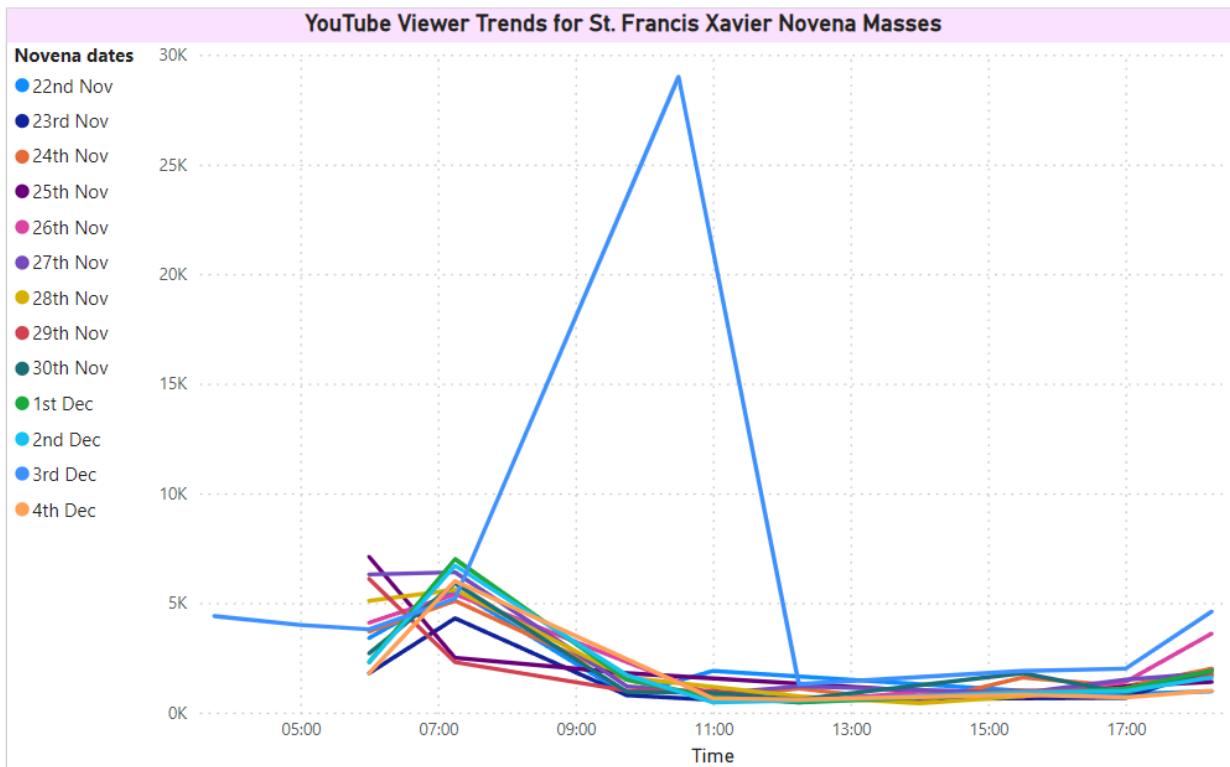


Fig.11. Viewer Trends for Novena Masses on the Official YouTube Channel

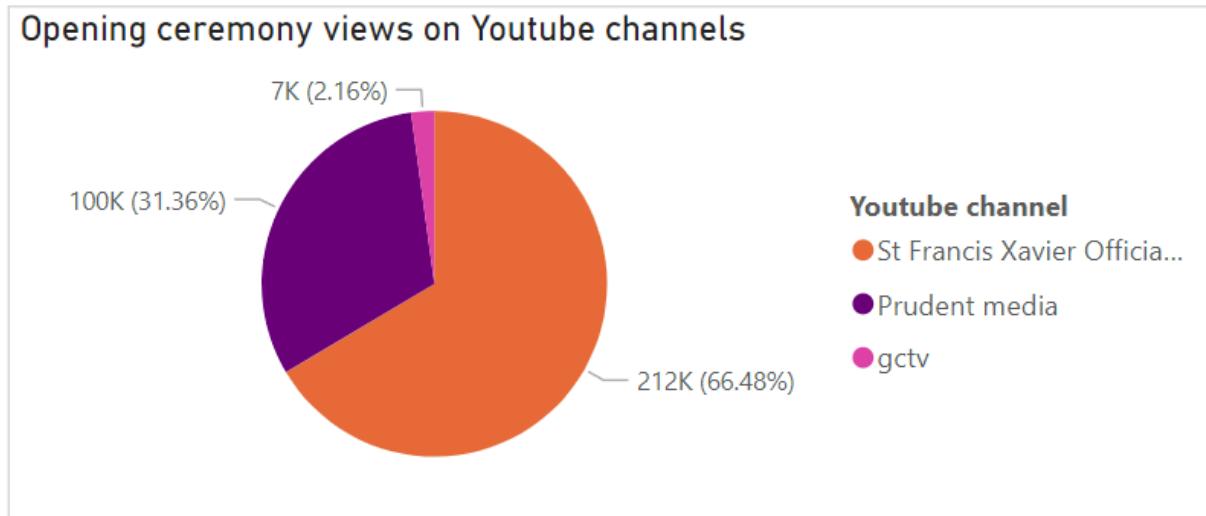


Fig.12. Viewership Across Various YouTube Channels for the Opening Ceremony

The viewer trends for masses were observed from the official YouTube channel of the Basilica of Bom Jesus, Old Goa, with the title 'St Francis Xavier Official Basilica Bom Jesus Goa,' <https://www.youtube.com/c/STFRANCISXAVIEROFFICIALBasilicaBomJesus> (Fig.11) covering all masses, including the language-specific ones, from the inaugural ceremony to the feast day on 3rd December. Approximately 212,700 people viewed the opening ceremony mass and procession of the Exposition on the official channel, while other media channels, such as Prudent and GCTV, recorded 100,000 and 6,900 views, respectively (Fig.12). The data reveals a significant spike in viewership around 10:30 AM on 3rd December, corresponding to the High Mass, indicating its heightened engagement. Early morning masses, between 6:00 AM and 7:15 AM, exhibited increased viewer counts, reflecting steady audience participation, while evening masses, such as those at 5:00 PM and 6:15 PM, showed a similar rise in engagement, especially for English masses that likely attracted a broader or international audience. Konkani masses consistently dominated most time slots, demonstrating a strong local or regional viewership, while masses in languages like Tamil, Portuguese, Kannada, and French highlighted the committees efforts to accommodate diverse pilgrim communities. Viewer counts varied across dates, with certain novena days, such as 28th November, witnessing higher engagement, possibly due to special ceremonies or enhanced promotional efforts on those days.

Year-wise YouTube Viewership Trends for the Exposition

An analysis of YouTube viewer counts across key Exposition days including Inaugural, and Feast Day reveals the evolving patterns of digital engagement. Notably, feast day viewership remained below 20,000 until 2019 but surged drastically in 2020 (86,185 views) and peaked in

2021 (138,400 views), reflecting pandemic-driven shifts to online participation. Although physical attendance resumed in 2022, virtual participation remained significant, with approximately 47,100 views in 2022 and 34,700 in 2023. In 2024, concurrent with the 18th Exposition, Feast Day viewership rose back to 56,200 despite the return of large in-person gatherings indicating the continued relevance of hybrid participation. The inaugural mass of 2024 alone drew 318,900 views, underscoring heightened global interest. These figures demonstrate that online platforms have become integral to religious tourism, complementing on-site experiences and enabling broader outreach.

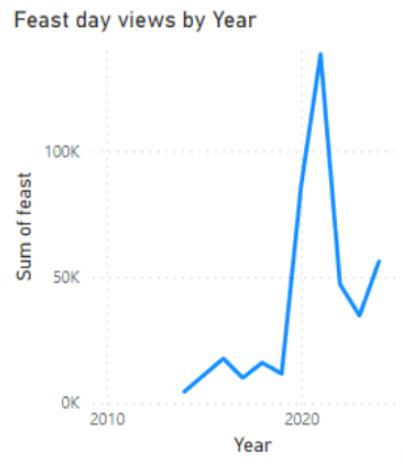


Fig.13.Feast Day YouTube viewership across years (2014–2024) showing a peak during the pandemic and sustained engagement thereafter.

3.3 Comparison Between Physical and Digital Participation

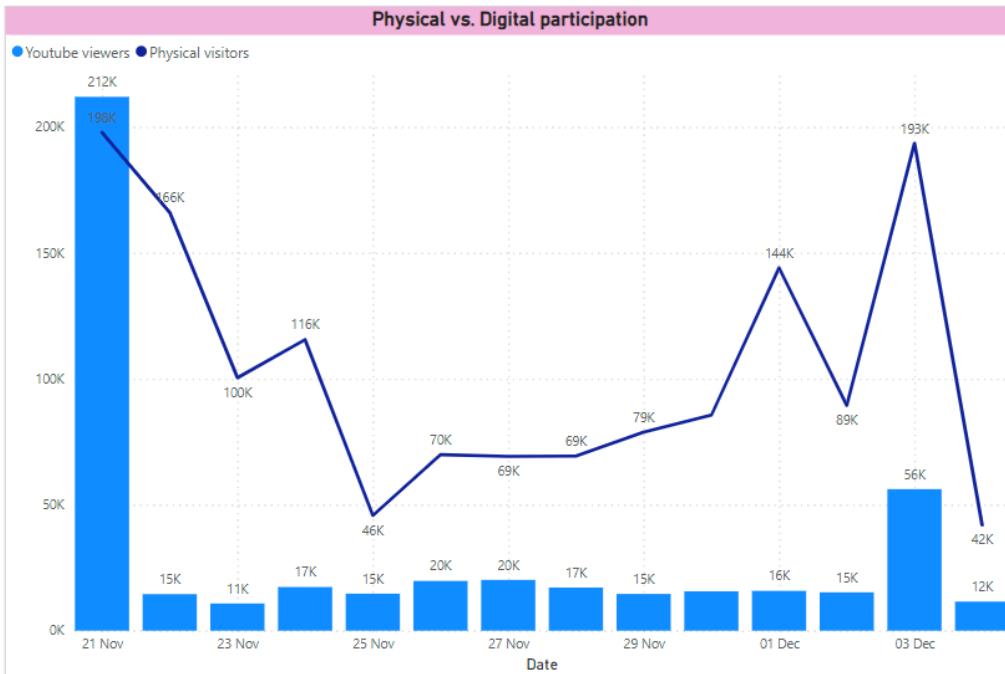


Fig.14. Participation of Viewers and Attendees Across Physical and Digital Platforms

The analysis of participation reveals distinct trends between physical visitors and digital viewers during the Exposition of the Sacred Relics of St. Francis Xavier. Physical attendance showcased steady numbers across the event, with notable peaks on significant days, such as the inaugural ceremony and the Feast Day on December 3rd, which drew the highest footfall at 193,494 visitors. Similarly, digital engagement through YouTube exhibited parallel patterns, with a major spike of 56,200 viewers on Feast Day. Daily YouTube viewership ranged between 10,795 and 20,170 during the Novena days, demonstrating consistent interest, while physical attendance varied significantly, reflecting the dynamics of on-site participation. The concurrent high engagement in physical and digital formats underscores the event's dual appeal, allowing it to reach both local attendees and a global audience effectively. This integration of traditional and virtual participation highlights the event's ability to adapt to contemporary needs while retaining its core religious significance.

3.4 Survey Response Overview

A total of 2,165 responses were collected on-site during the Novena days and from visitors staying in pilgrim accommodations, with an additional 115 responses gathered through an online survey form (Fig.15). Both genders participated equally in the survey shown in Fig.16. The summary statistics of the survey are presented in Table 4.1.

Table 4.1. Demographic and Motivational profile of Survey Respondents

Category	Variable	Percentage
Gender	Male	51.3
	Female	48.7
Age Group	Early professionals (25-34 years)	24.4
	Mature Adults (45-64 years)	23.5
	Young Adults (35-44 years)	18
	Youth (18-25 years)	15.9
Region	Young Dependents (below 18 years)	10.8
	Senior citizen (65+ years)	7.5
Region	South India	36.6
	West India	24
	Goa	18.3
	North India	6.8
	East India	6.2
	Europe	3.7
	Northeast India	1.7
	North America	1.3
	Asia	0.8
Visit Frequency/Pilgrim type	Oceania	0.6
	Africa	0.1
	First-time visitors	60.8
	Long-Time Devotees (10+ years)	19.1
	Occasional visitors (2-5 years)	17.8

	Frequent visitors (6-9 years)	2.3
Motivations/ Event	Veneration of Relics	39.45
	Mass	28.02
	Fair/Cultural Event/Sightseeing	32.53

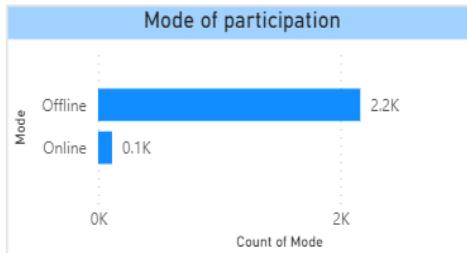


Fig.15. Mode of survey participation

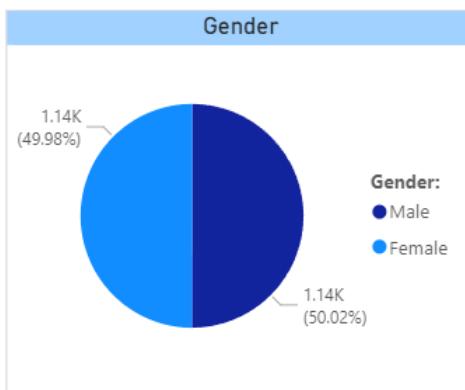


Fig.16. Gender of survey participants

The diverse age distribution of participants highlights the demographic appeal of the event. From Fig.17 we see that the majority of attendees were from working age groups, that includes early professionals (25–34 years) and mature adults (45–64 years). This indicates its relevance to the individuals who like balancing their professional and personal commitments. The youth (18–25 years) and young adults (35–44 years), participated in huge numbers too. Family participation was evident through the young dependents (below 18 years), while the presence of senior citizens (65+ years) showed the spiritual connection and devotion of the older generations. This multigenerational participation emphasized the broad appeal of the Exposition, fostering a gathering that bridged age groups through shared faith and tradition.

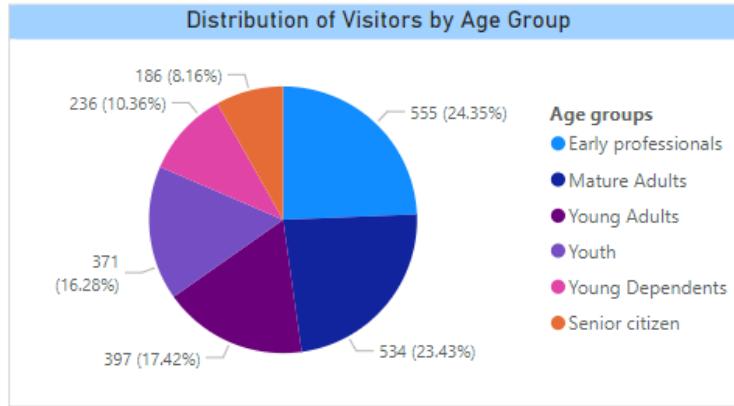


Fig.17.Distribution of visitors that took part in the survey by their age

Geographical Analysis of Visitors

The Basilica of Bom Jesus attracted visitors from many places that included Asia, North America, Oceania, and Europe, including Germany, the UK, and Ireland. This proves its universal importance and its status as a UNESCO World Heritage Site. Goa, the host state, had the largest share of visitors, as seen in Fig.19. Followed by Maharashtra, Kerala, Karnataka, and Tamil Nadu which also showed immense participation. Their proximity and accessibility to Goa played an important role in the high footfall. Visitors from across India, including Uttar Pradesh, Odisha, and Rajasthan, highlight its past connections and faith. A closer look at the distribution in India reveals that most of the visitors came from South India, followed by West India and Goa and thus highlights its connection in these parts of the country. While North and East India also contributed to the visitors representation, theirs was comparatively moderate (Fig.20). Lower participation rates from Asia, Oceania, and Africa suggested the untapped potential for targeted outreach in these regions.

This analysis has shown us the Basilica's ability to attract a broad array of visitors, with strong engagement regionally and with an international footprint that needs expansion. To further enhance its global reach and inclusivity, strong initiatives need to be taken to engage the underrepresented regions, which would lead to higher participation from not only domestic but also international audiences.



Fig.18. Geographical distribution of visitors

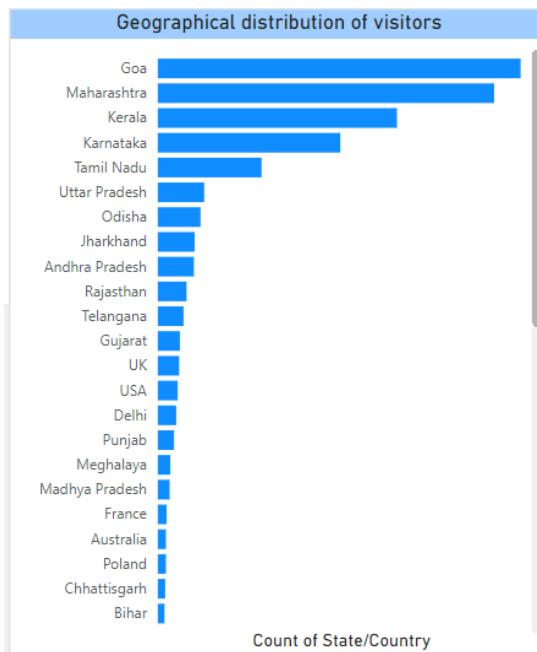


Fig.19. Goa, as the host state, shows the highest participation of visitors

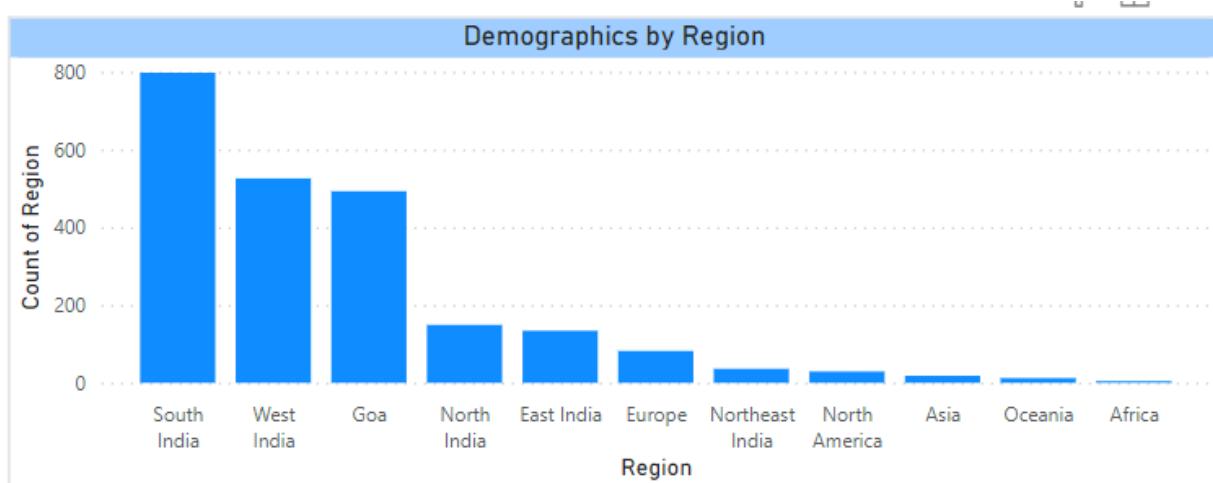


Fig.20. Region Wise geographical distribution of visitors

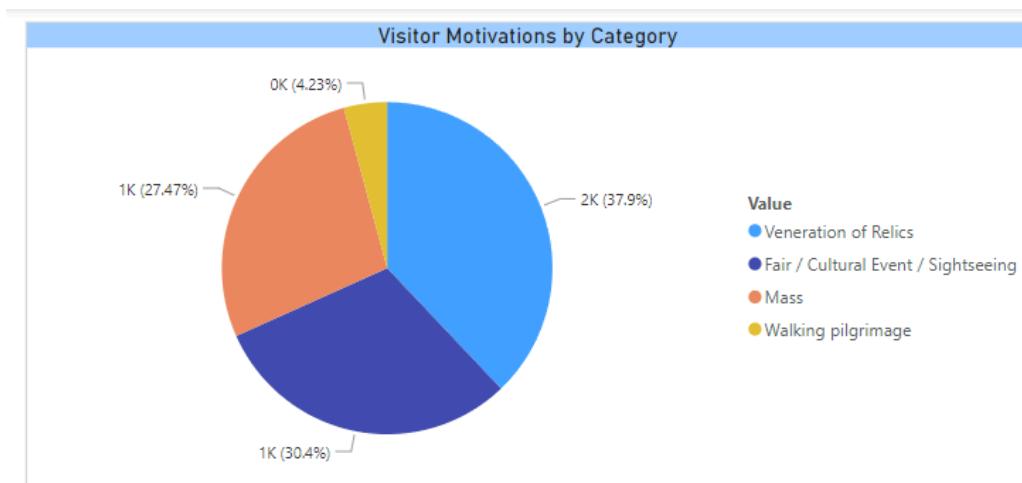


Fig.21. Motivations of participants behind visiting the Basilica of Bom Jesus

There is a blend of motivations behind visiting the Basilica of Bom Jesus that includes spiritual devotion, cultural engagement, and traditional practices, that reflects its multifaceted appeal. Fig.21 shows that a significant portion of visitors (37.9%) are drawn by the veneration of relics, highlighting deep reverence for St. Francis Xavier, which symbolizes their deep faith. Similarly, Mass attendance (27.4%) shows their religious practices, with recurring visits for spiritual gatherings and special services reflecting the Basilica's active role as a place of worship. The cultural events, fairs, and sightseeing attract 30.4% of visitors, illustrating the site's broader appeal as a hub for Goan tourism and community engagement. This part of visitors value the Basilica's architectural majesty, its historical significance, and associated festivities, exhibiting the intersection of religion and tourism. Walking pilgrimages, though representing a smaller

share (4.23%), show that some visitors also have a deep personal respect and spiritual practice with penance and devotion. The Basilica's ability to cater to diverse visitor interests is seen here, balancing its spiritual essence along with its cultural and historical prominence. This visual analysis underscores the evolving dynamics of religious tourism and shows us the need to sustain both its spiritual sanctity and cultural heritage.

The overnight walking pilgrims shared that their motivations extended beyond mere participation in religious events, such as attending the exposition or Mass. Their pilgrimage was an act of repentance, devotion for the saint, and praise to God.

These pilgrims covered 130-140 kilometers for over 2-4 days on foot, singing hymns, reciting the rosary, praying and attending daily Mass along the way. It is said that this tradition of walking pilgrimage began in 1974 with just seven individuals and has grown steadily since then. Many participants described their feelings as a deeply fulfilling experience and a ritual passed down through generations.

The pilgrims arrived at the venue early in the morning on December 2nd and made their return journey after attending the 3:00 AM Mass on the Feast Day. To accommodate their needs, camping facilities were provided, which enhanced their overall experience and supported this enduring tradition.

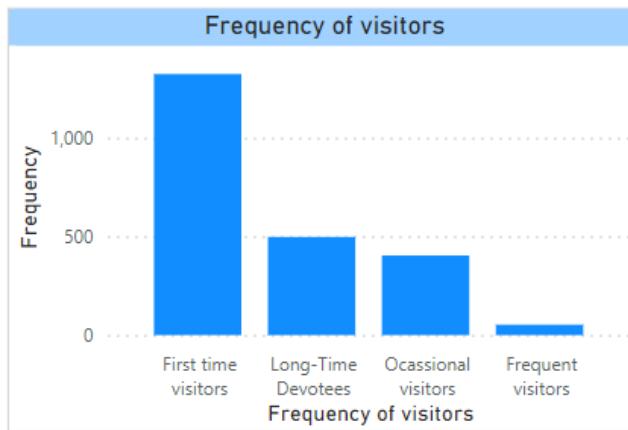


Fig.22. Frequency of visitors

The frequency of visits in Fig.22 indicates diverse patterns of visitors, offering insights into their devotion for the saint. Major portion of visitors have participated for the first-time, showcasing the curiosity that the Basilica brings and its ability to attract new audiences. Long-time devotees, visiting for over 10 years, form a dedicated group who maintain a deep spirituality towards the site, often returning out of reverence and tradition. Occasional visitors, attending between 2 to 5 years, represent a mix of cultural interest and religious engagement, while frequent visitors,

returning between 6 to 9 years, suggest a group that balances consistent devotion with a more flexible visiting pattern. These varied visiting frequencies underline the Basilica's dual role as a center for enduring faith and an accessible destination for new and occasional visitors.

The accommodation choices of pilgrims reveal significant differences in preferences and logistical considerations. While few pilgrims opted for the on-site camps and barracks, the larger group of non-campers demonstrated a greater demand for more private or comfortable lodging options. The campers, including Goan and Indian pilgrims as well as walking pilgrims, might have chosen the camps for their proximity to the Basilica, communal experience, or budget-friendly nature.

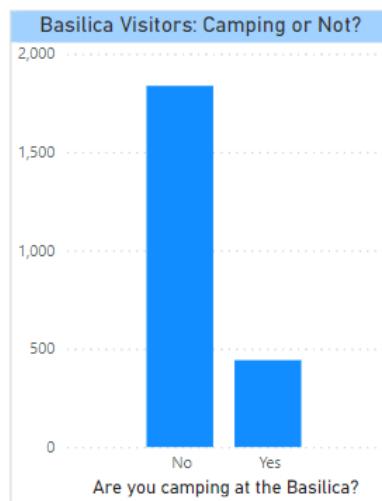


Fig.23.Participants camping at the Basilica

3.5 Goa Pilgrim Camp Check-in Analysis

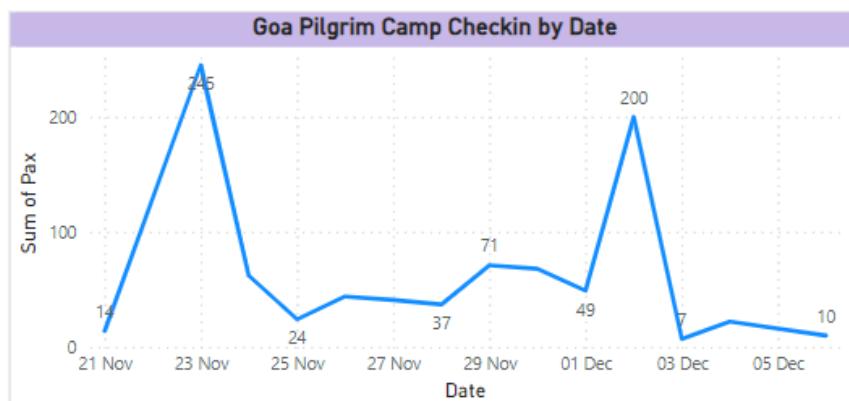


Fig.24. Goa pilgrims camp check-ins

The analysis of overnight pilgrims in Fig.24 reveals key trends in check-ins at the Pilgrim Village and camping facilities during the Exposition. Peak check-ins were observed on 22nd November 2024, just before the Novenas began, and again on 2nd December, aligning with the Feast of St. Francis Xavier on 3rd December. This pattern underscores the heightened influx of pilgrims during these significant events. A steady increase in check-ins was noted as the Novenas progressed, reflecting the growing participation and anticipation leading up to the feast day.

3.6 Pilgrim Village Check-in Analysis

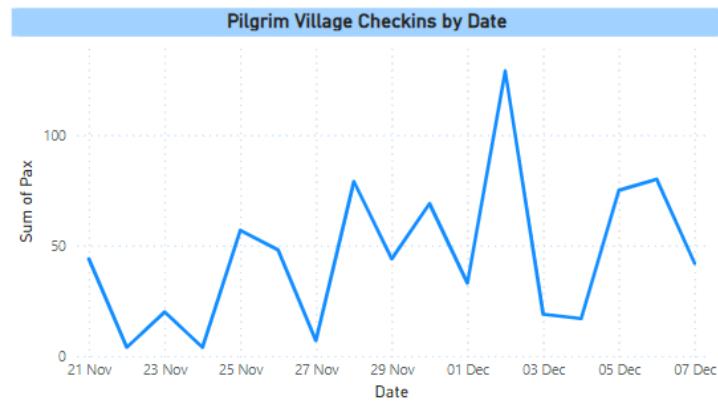


Fig.25. Pilgrim Village check-ins by date

In Fig.25. we observe the fluctuations in the daily check-ins at the Pilgrim Village accommodations during the Exposition. Check-in peaks were observed around 24th and 25th November marking the first days of the novenas, followed by a huge surge on 1st and 2nd December before the feast day, and then 5th December a weekend, suggesting higher footfall during weekends since pilgrims came to attend Sunday masses. Guests typically stayed for 2 to 3 days, indicating their short stay.

The Pilgrim Village had many lodging options, including camps and barracks, to accommodate the walking pilgrims and other visitors. Feedback was taken from the pilgrims about their stay and a positive experience was observed,since staying nearby provided easy access to the Exposition site and other related religious and cultural activities.

Apart from this, there were around 1,200 walking pilgrims who had traveled from various parts of India, including Maharashtra (Sindhudurg, Pune, Kolhapur, Dodamarg, and Sawantwadi) and Karnataka (Belgaum) preferred to stay in tents that were specially built for them, highlighting the need for convenient and budget friendly accommodations them near the venue.

3.7 Correlation Analysis

The initial step involved analyzing associations between continuous or ordinal variables, such as Age and Number of Visits. To assess the strength and direction of relationships among numerical variables Pearson correlation is used. A correlation heatmap was generated to visualize associations. Notable correlations included Age and Number of Visits: 0.38 which shows a positive correlation between the two, which shows that elder pilgrims visit the Basilica more often.

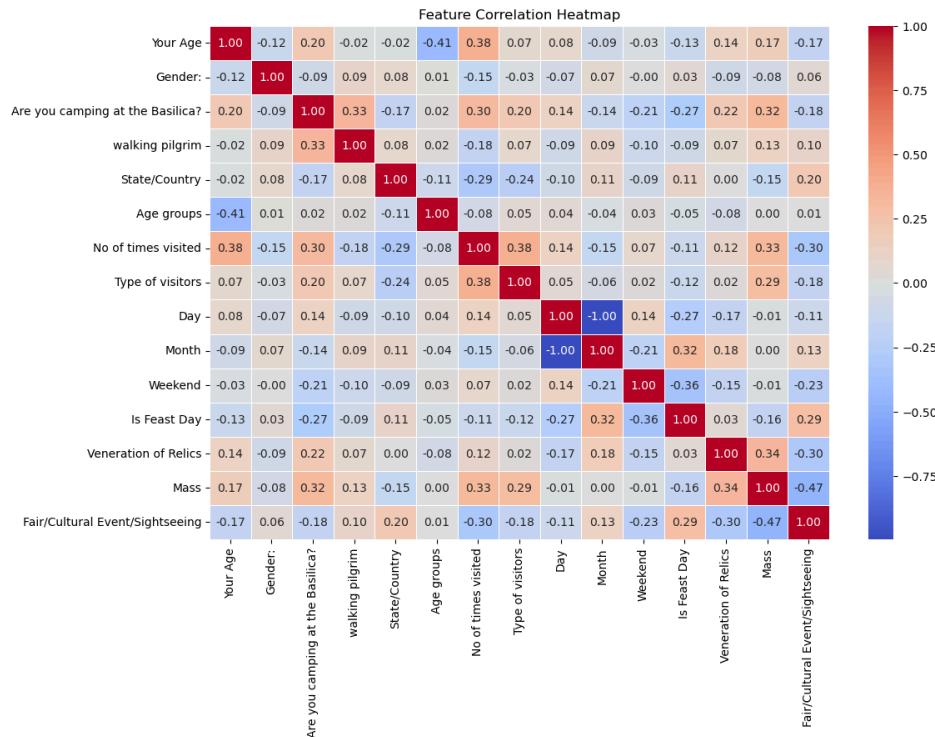


Fig.26. Correlation Strength (Pearson's Correlation) between features

3.8 Association Analysis

A series of chi-squared tests of independence were conducted. The predictors included demographic variables (e.g., age group, gender, state/country), visit-related variables (e.g., type of visitor, camping status), and calendar-related variables (e.g., weekend, feast day). The dependent variables were binary indicators of participation in the events (Mass, Veneration of Relics, Fair/Cultural Events/Sightseeing)

The strength of association for each significant predictor was further evaluated using Cramér's V, where values closer to 1 indicate stronger associations.

Table 4.2 Association strengths between the predictor variables and target events

Mass		
Predictor	Chi² p-value	Cramér's V
Type of visitors	0	0.3974
State/Country	0	0.3448
Are you camping at the Basilica?	0	0.3211
Weekend	0	0.2283
Is Feast Day	0	0.1561
Gender:	0.0001	0.082
Age groups	0.1398	0.0234
Fair/ Cultural Event/ Sightseeing		
Predictor	Chi² p-value	Cramér's V
State/Country	0	0.3724
Is Feast Day	0	0.2855
Type of visitors	0	0.2846
Age groups	0	0.264
Are you camping at the Basilica?	0	0.1824
Gender:	0.0005	0.0723
Weekend	0.5232	0
Veneration of Relics		
Predictor	Chi² p-value	Cramér's V

State/Country	0	0.3466
Are you camping at the Basilica?	0	0.2149
Weekend	0	0.1518
Age groups	0	0.1392
Type of visitors	0	0.1287
Gender:	0.0116	0.0498
Is Feast Day	0	0

Mass Participation

- Type of Visitor exhibited the strongest association (Cramér's V = 0.3974), indicating that repeat or frequent visitors are more likely to attend Mass.
- State/Country (V = 0.3448) and Camping at the Basilica (V = 0.3211) also showed notable associations.
- Temporal factors such as Weekend (V = 0.2283) and Feast Day (V = 0.1561) had weaker but statistically significant associations.
- Gender had a minimal effect (V = 0.082), indicating a more uniform participation across male and female visitors.

Veneration of Relics

- State/Country (V = 0.3466) and Camping at the Basilica (V = 0.2149) were the top predictors.
- Age Group and Type of Visitor also demonstrated small to moderate associations (V = 0.1392 and 0.1287, respectively).
- Gender and calendar-based variables had negligible associations (Cramér's V < 0.05).

Fair/Cultural Events and Sightseeing

- State/Country was again the most influential predictor (V = 0.3724), followed by Feast Day (V = 0.2855) and Type of Visitor (V = 0.2846).
- Age Group (V = 0.264) had a moderately strong relationship, suggesting younger or middle-aged visitors may be more drawn to cultural aspects.
- Predictors such as Camping, Gender, and Weekend had lower or negligible associations (V < 0.18).

Implications for Prediction

The features with the highest Cramér's V scores are strong candidates for inclusion in classification models to predict participation in specific religious events. E.g., Predicting Mass attendance can effectively use: Type of Visitors, State/Country, Camping, and Age Groups.

Predicting Cultural Event attendance can benefit from using: State/Country, Feast Day, Type of Visitor.

Interpretation of Results

The Fig.27 plot visualizes Cramér's V values, showing the strength of association between categorical features and the type of event attended. Key insights include:

- Mass attendance has the strongest association with Type of visitors (whether they are repeat or first-time pilgrims), as well as geographic ($V = 0.345$), and whether they are camping at the Basilica ($V = 0.321$).
- Fair/Cultural Event/Sightseeing shows notable associations with State/Country ($V = 0.372$) and Type of visitors ($V = 0.285$).
- Veneration of Relics is also influenced by State/Country ($V = 0.347$), followed by Camping and Age groups.

Overall, visitor type, location, and pilgrimage behavior (like camping or overnight stay) emerged as influential factors in predicting event participation patterns.

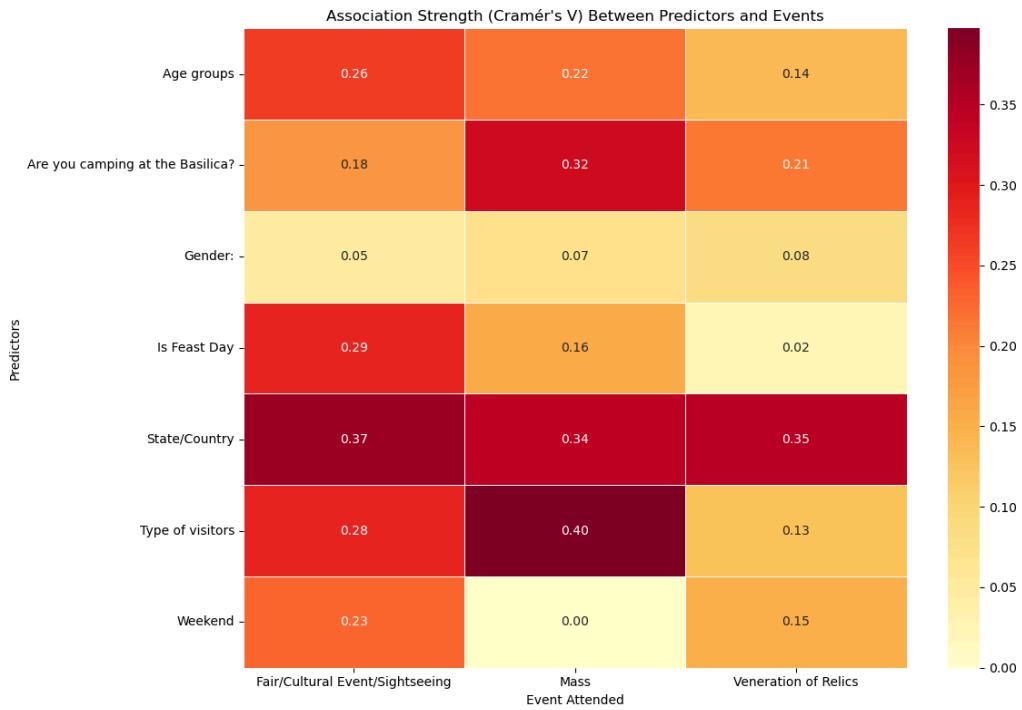


Fig.27. Association Strength (Cramer's V) Between Predictors and Events

3.4 Hypothesis Results

H1: Demographic factors (age, gender, region) significantly influence the likelihood of participating in the Exposition.

Chi-Square Tests (Age Group vs. Event Attendance):
 Veneration of Relics: Chi² = 46.9349, p-value = 0.0000
 → There is a significant relationship between age group and attendance at Veneration of Relics.
 Mass: Chi² = 108.3130, p-value = 0.0000
 → There is a significant relationship between age group and attendance at Mass.
 Fair/Cultural Event/Sightseeing: Chi² = 155.8001, p-value = 0.0000
 → There is a significant relationship between age group and attendance at Fair/Cultural Event/Sightseeing.

Fig.28. Chi Square Test results for relationship between Age Groups and Event Attendance

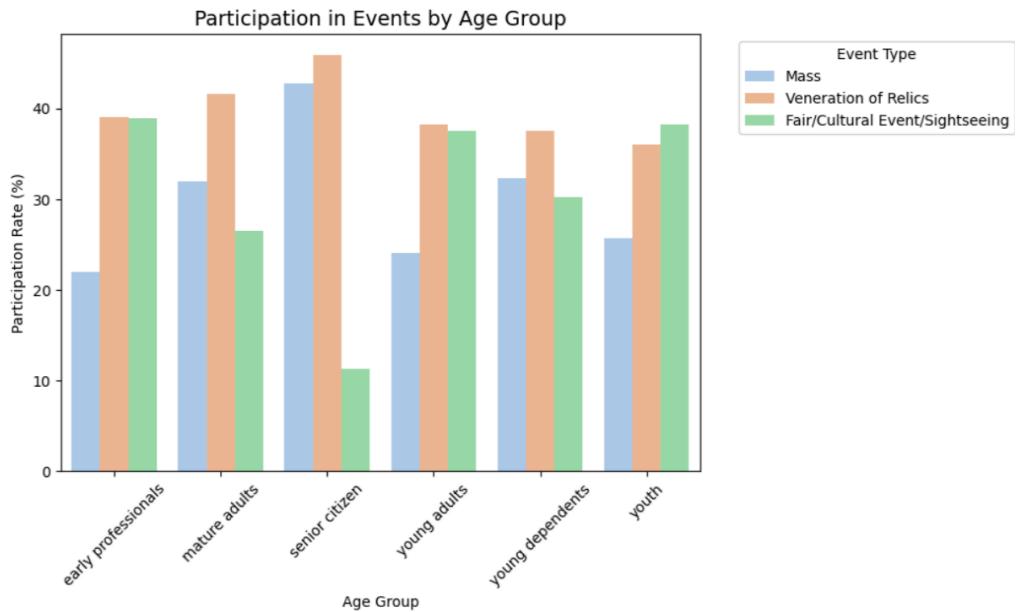


Fig.29.Age-wise engagement in Pilgrimage activities

Fig.29 presents a breakdown of age-wise engagement in three key pilgrimage-related activities. Older age groups, especially senior citizens and mature adults, show a much higher engagement in religious rituals, with 42.8% and 32% participation in Mass, and 45.9% and 41.6% in relic veneration respectively. In contrast, the younger pilgrims, particularly youth, young adults, and early professionals, show a strong inclination toward cultural tourism, with participation rates more than 37% in that category. It is noticed that senior citizens do not engage much in cultural events (11.3%), further proving the trend of age-based preference for spiritual versus socio-cultural experiences. This pattern indicates that pilgrimage engagement is shaped by age, with younger attendees seeking broader experiential elements, while older pilgrims focus more on traditional devotional practices.

Chi-Square Tests (Gender vs. Event Attendance):
Veneration of Relics: Chi₂ = 15.5282, p-value = 0.0001
 → There is a significant relationship between Gender and attendance at Veneration of Relics.
Mass: Chi₂ = 12.2923, p-value = 0.0005
 → There is a significant relationship between Gender and attendance at Mass.
Fair/Cultural Event/Sightseeing: Chi₂ = 6.3661, p-value = 0.0116
 → There is a significant relationship between Gender and attendance at Fair/Cultural Event/Sightseeing.

Fig.30. Chi Square Test results for the relationship between Gender and Event Attendance

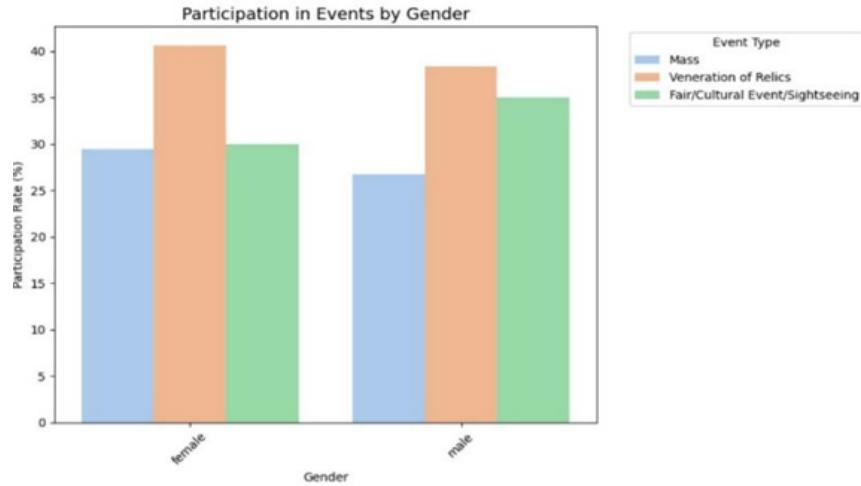


Fig.31. Gender-wise engagement in Pilgrimage activities

Fig.31 indicates that although engagement levels are generally similar, minor trends emerge, with females slightly more inclined towards Mass and Veneration, while males show a bit more interest in cultural events.

Chi-Square Tests (Region vs. Event Attendance):
 Veneration of Relics: Chi₂ = 67.4211, p-value = 0.0000
 → There is a significant relationship between Region and attendance at Veneration of Relics.
 Mass: Chi₂ = 218.9967, p-value = 0.0000
 → There is a significant relationship between Region and attendance at Mass.
 Fair/Cultural Event/Sightseeing: Chi₂ = 256.6424, p-value = 0.0000
 → There is a significant relationship between Region and attendance at Fair/Cultural Event/Sightseeing.

Fig.32. Chi Square Test results for relationship between Region and Event Attendance

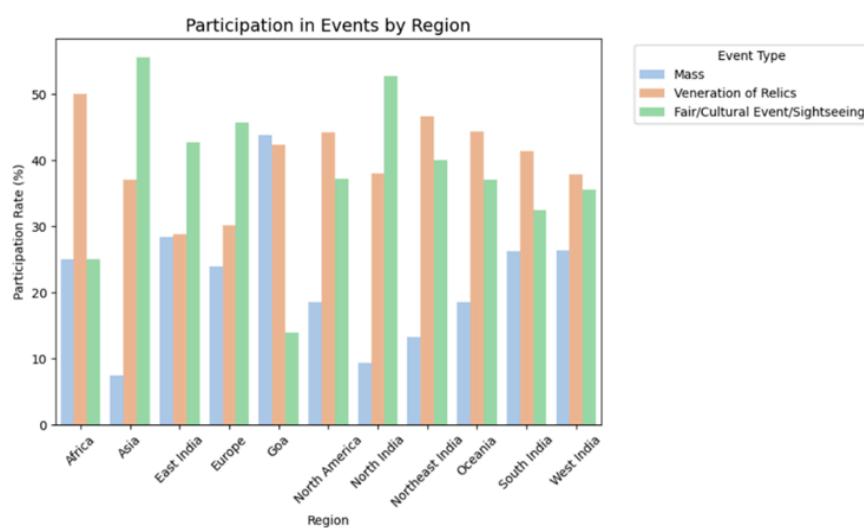


Fig.33. Region-wise engagement in Pilgrimage activities

		Mass	Veneration of Relics	Fair/Cultural Event/Sightseeing	
Region					
Africa	1		2		1
Asia	2		10		15
East India	67		68		101
Europe	31		39		59
Goa	299		289		95
North America	8		19		16
North India	20		82		114
Northeast India	8		28		24
Oceania	5		12		10
South India	366		579		453
West India	256		367		345

		Mass	Veneration of Relics	Fair/Cultural Event/Sightseeing	
Region					
Africa	25.0		50.0		25.0
Asia	7.4		37.0		55.6
East India	28.4		28.8		42.8
Europe	24.0		30.2		45.7
Goa	43.8		42.3		13.9
North America	18.6		44.2		37.2
North India	9.3		38.0		52.8
Northeast India	13.3		46.7		40.0
Oceania	18.5		44.4		37.0
South India	26.2		41.4		32.4
West India	26.4		37.9		35.6

Fig.34. Region-wise engagement contingency table

Fig.33. bar graph shows the breakdown of event participation based on the visitor's region. Notable differences in Mass and Cultural Event attendance are observed across states and countries, indicating that geographical and cultural contexts shape visitor motivations.

H2: First-time visitors have different engagement patterns compared to long-time devotees and frequent pilgrims.

```
-- H2: Mass Engagement by First-Time vs Repeat Visitors --
Chi-Square Statistic: 268.2691021010068
p-value: 2.7049255715967985e-60
Significant difference in Mass engagement patterns between first-time and repeat visitors.

-- H1: Veneration of Relics Engagement by First-Time vs Repeat Visitors --
Chi-Square Statistic: 5.520804920718159
p-value: 0.018791610339722386
Significant difference in Veneration of Relics engagement patterns between first-time and repeat visitors.

-- H1: Fair/Cultural Event/Sightseeing Engagement by First-Time vs Repeat Visitors --
Chi-Square Statistic: 109.53497932655658
p-value: 1.238987339237257e-25
Significant difference in Fair/Cultural Event/Sightseeing engagement patterns between first-time and repeat visitors.
```

Fig.35. Chi-Square Test Results -Engagement Patterns in First-Time vs Repeat Visitors

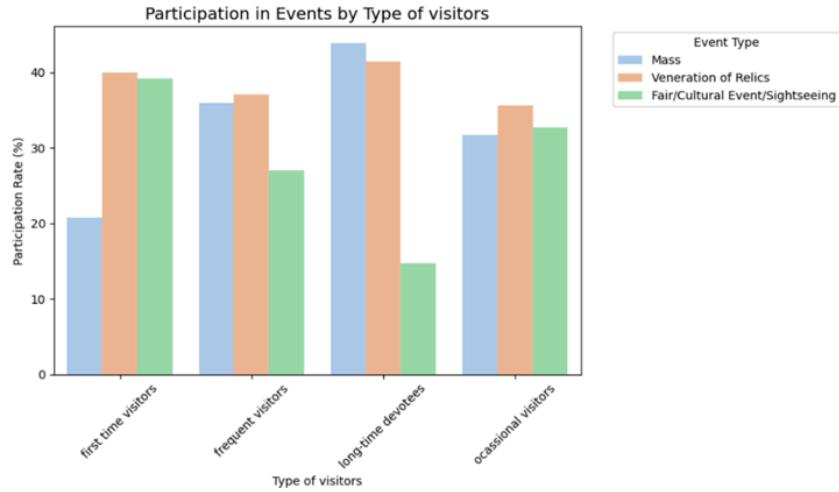


Fig.36. Visitor Engagement Trends Based on Visit Frequency

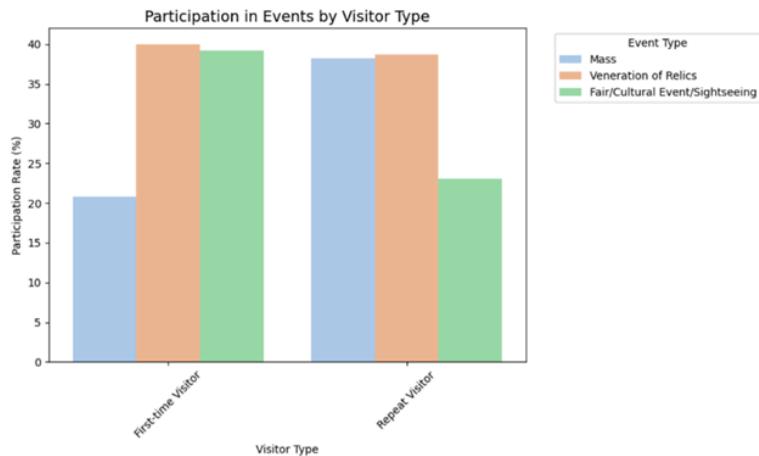


Fig.37. Visitor Engagement in first-time and repeat pilgrims

Count Table:			
Visitor Category	Mass	Veneration of Relics	Fair/Cultural Event/Sightseeing
First-time Visitor	460		884
Repeat Visitor	603		611
Percentage Table:			
Visitor Category	Mass	Veneration of Relics	Fair/Cultural Event/Sightseeing
First-time Visitor	20.8	40.0	39.2
Repeat Visitor	38.2	38.7	23.1

Fig.38. Contingency table depicting % of event engagement in first-time and repeat pilgrims

Fig.38. depicts that first-time visitors demonstrated a higher participation rate in Fair/Cultural Events/Sightseeing (39.2% of their total participation), indicating a more touristic or exploratory approach to their visit.

They participated in Mass (20.8%) and Veneration of Relics (40.0%) to a lesser extent, suggesting that religious practices are secondary to their cultural and sightseeing interests.

Repeat visitors (which includes frequent visitors, long-time devotees, and occasional visitors) showed notably higher participation in Mass (38.2%) and Veneration of Relics (38.7%), indicating a stronger spiritual connection and more established pilgrimage habits.

Repeat visitors exhibited lower participation in Fair/Cultural Events/Sightseeing (23.1%), suggesting a focus on the religious aspects of the pilgrimage rather than cultural or sightseeing experiences.

H3: The Exposition had no significant effect on the total number of tourist arrivals in 2024 compared to previous years

To assess whether the 2024 Exposition had really contributed to the observed surge in tourist arrivals in Goa as seen in Appendix II, a hypothesis test was conducted comparing tourist figures from 2024 (up to December) to the previous year (2023). Using a two-sample t-test, the results indicated a statistically significant increase in tourist arrivals in 2024 ($p\text{-value} < 0.05$), supporting the hypothesis that the Exposition did have a measurable impact on tourism in Goa.

Z-score: inf

P-value: 0.0000

Significant increase: 2024 tourist arrivals are significantly higher than 2023.

Fig.39.Two-sample Z-test to check the increase in tourist arrivals in 2024 compared to 2023

The Z-score, reported as infinite, was most likely influenced by the extremely small standard error, which occurs when comparing single-year totals without monthly-level variation, thereby reducing variance. This highlights the magnitude of the difference in the yearly totals and strengthens the statistical significance of the finding.

A Times of India article also reported a marked increase in Goa's tourist footfall by the end of the Exposition, further corroborating this result ([TOI, 2024](#)). This aligns with the historical role of the Exposition in attracting both domestic and foreign pilgrims, validating its influence beyond purely religious dimensions and underscoring its broader contribution to Goa's tourism-driven economy.

This result reinforces the visual and anecdotal evidence presented earlier, 2024 marked a historic surge in tourism in Goa, largely attributed to this special event, the 18th Exposition of the Relics of St. Francis Xavier.

4.5 Model Evaluation Results

To predict the engagement patterns of visitors during the Exposition, multiple supervised machine learning models were implemented and evaluated. The goal was to classify or predict the likelihood of participation in various event activities (e.g., Mass, Veneration of Relics, Cultural Events) based on demographic and behavioral features such as age group, type of visitor, state/country, and past attendance.

Gradient Boosting

--- Gradient Boosting Classifier ---				
	precision	recall	f1-score	support
Fair/Cultural Event/Sightseeing	0.71	0.86	0.78	369
Veneration of Relics	0.79	0.93	0.85	472
Mass	0.77	0.70	0.73	316
micro avg	0.76	0.84	0.80	1157
macro avg	0.76	0.83	0.79	1157
weighted avg	0.76	0.84	0.80	1157
samples avg	0.75	0.84	0.77	1157
 Accuracy Score: 0.4969230769230769				
Hamming Loss: 0.2528205128205128				
Jaccard Score (macro): 0.6524444647874068				
AUC-ROC Score (macro): 0.792014695644499				

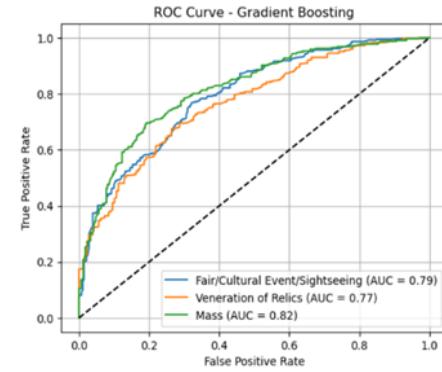


Fig.40. Classification report and ROC curve for Gradient Boosting model

Gradient Boosting stands out as the best-performing model overall, with the highest accuracy (49.7%), F1-Score (0.77), and Jaccard Score (0.65). It strikes a good balance between precision and recall, ensuring consistent performance across classes, especially for more complex patterns. Its low Hamming Loss (0.25) indicates fewer incorrect label predictions, making it a robust choice for multi-label classification tasks. Jaccard Score shows more overlap between predicted and true labels than Random forest (0.608). The samples average F1 score indicates stronger performance per sample. A recall of 0.93 for Veneration of Relics is especially high, meaning this model rarely misses this label. Mass has slightly lower recall (0.70), but maintains a high AUC (0.82), suggesting the model is generally good at ranking but may be threshold-sensitive. Overall the gradient boosting will have fewer label-wise mistakes and greater ability to distinguish classes. One downside is that Gradient Boosting can be slower to train, especially on larger datasets, which may require tuning to optimize performance. Despite this, it remains the top performer for your use case.

Random Forest

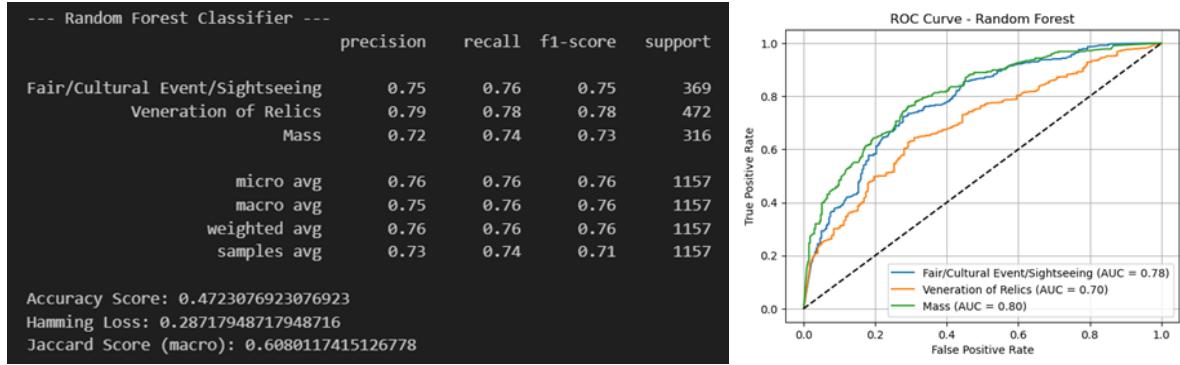


Fig.41. Classification report and ROC curve for Random Forest model

Random Forest offers a balanced performance across all metrics, with an accuracy of 47%, which is not very ideal for multi-label problems but it is lower because accuracy considers all labels to be predicted correctly. An F1-Score of 0.75 on an average indicates balanced performance. Precision of 0.75 to 0.79 indicates low false positives across all labels which is very good. Recall of 0.74 to 0.78 shows that the false negatives are low too. The model is capturing the relevant events. F1 score of 0.75 shows balanced performance. Its slightly higher Hamming Loss (0.29) reflects a few more incorrect predictions compared to Gradient Boosting. Jaccard score of 0.60 is decent for multi-label problems. AUC values suggest reliable discrimination for most events. It provides an interpretable and versatile approach, capturing a wide range of patterns in the data. Its ability to generalize well across different event types makes it a reliable baseline model. While Random Forest is good overall, it may not perform as well in edge cases or with very specific patterns compared to models like Gradient Boosting.

k-Nearest Neighbors (KNN)

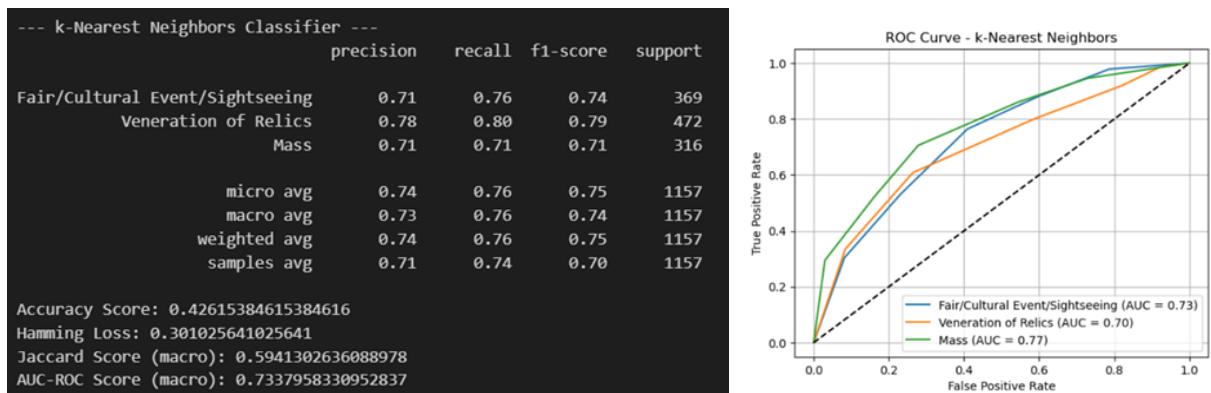


Fig.42. Classification report and ROC curve for k-NN model

KNN provides consistent performance across all classes, with an accuracy of 43%, a Jaccard Score of 0.60, and an F1-Score of 0.71. It works well when proximity-based patterns are strong in the data. Jaccard Score and AUC-ROC are lower, showing reduced ability to distinguish and correctly overlap labels. This model is effective in situations where events are closely related and follow similar patterns. The main drawback of KNN is scalability. It can be computationally expensive and slow with larger datasets, especially when the dataset grows significantly. This limits its practicality for very large-scale tasks, though it remains a solid choice for moderate-sized datasets.

Decision Tree

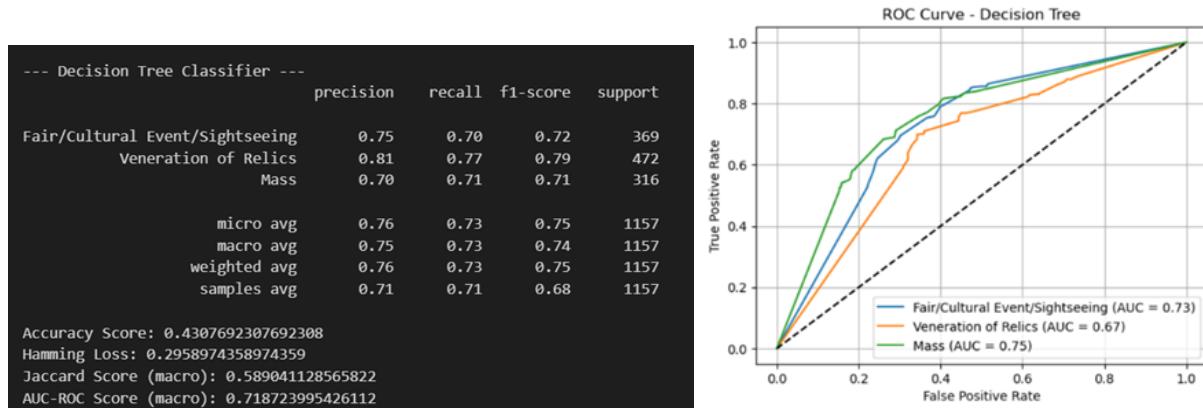


Fig.43. Classification report and ROC curve for Decision tree

Decision Trees are simple to interpret and easy to explain, which is a major strength when interpretability is a priority. It offers a good, straightforward solution. They tend to capture patterns well in relatively small or structured datasets. Hamming Loss and Jaccard Score suggest the model makes frequent, slightly imprecise predictions. One notable downside is their tendency to overfit, especially if not properly tuned which explains modest recall and lower AUC in a multi-label context. The model's performance is slightly weaker than Random Forest, as it struggles to generalize as effectively, reflected in the slightly lower metrics.

Logistic Regression

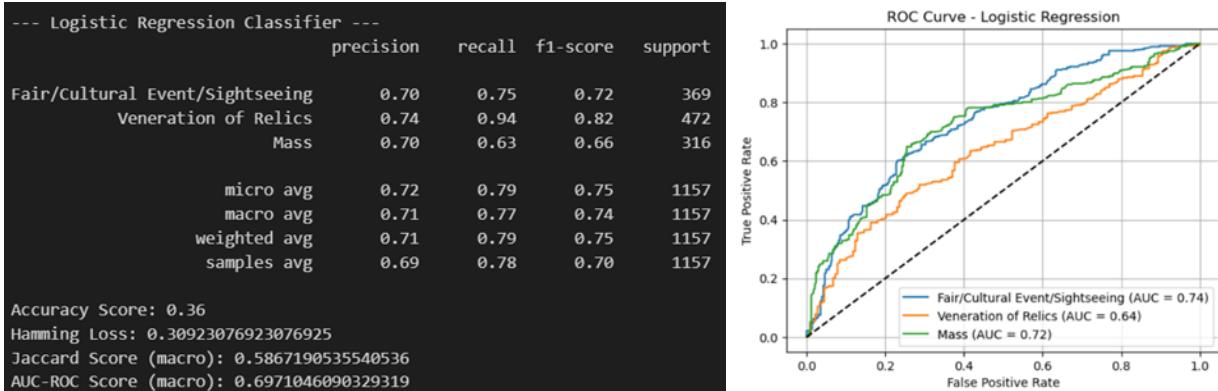


Fig.44. Classification report and ROC curve for Logistic Regression

Logistic Regression is known for its simplicity and speed. It is easy to implement and can work quickly with smaller datasets. It struggles with non-linear decision boundaries, which is likely why it underperforms in this multi-label scenario. It does show a high recall for specific event types like the "Veneration of Relics" class. Its major limitation is that it fails to capture complex relationships in multi-label tasks, making it less suitable for cases with more intricate patterns and overlapping features. The low accuracy (36%) and occasional complete miss of predictions (especially for the "Mass" event) are indicative of its inability to handle complex multi-label scenarios effectively. AUC-ROC Score of 0.697 is significantly lower which suggests weaker generalization power.

Support Vector Machine (SVM)

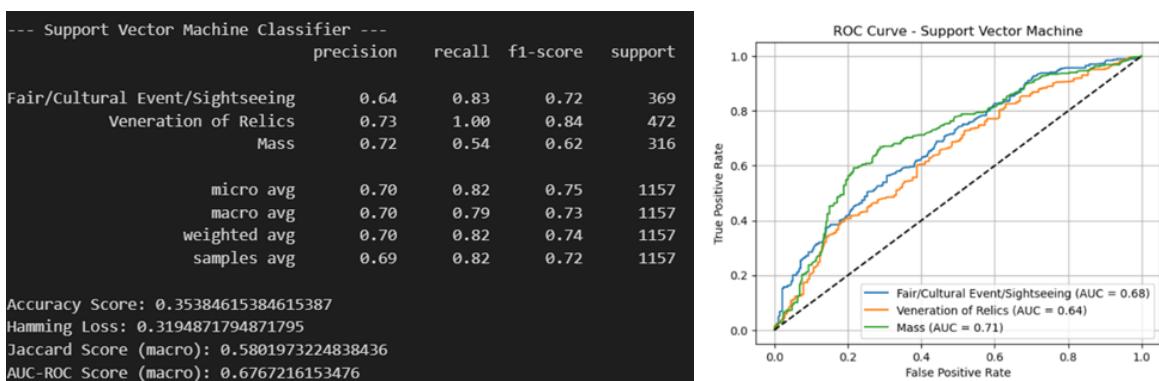


Fig.45. Classification report and ROC curve for SVM

SVM excels in recall for specific classes, particularly for the "Veneration of Relics" event, where it achieves a perfect recall score of 1.00 meaning the model catches all true instances of this label, though it may include some false positives. It also achieves a high F1-Score of 0.72 and performs decently on the overall classification task, making it suitable for applications where

recall is critical for certain event types. Fair/Cultural Event/Sightseeing has high recall (0.83) but lower precision (0.64), meaning it's good at identifying all events, but may mislabel others. Mass has weaker recall (0.54), suggesting many actual labels are missed. SVM's main drawback is its tendency to miss predictions entirely, especially for certain events like "Mass," leading to low accuracy (35%). This suggests that the model may not handle the full range of event types as effectively as other models. Moreover, its sensitivity to outliers and the potential for no predictions in some instances can make it unreliable for a more balanced prediction.

3.6 Model Performance Summary

The evaluation results are summarized in Table 4.3, demonstrating that Gradient boosting achieved the highest predictive performance, indicating its suitability for capturing complex patterns in visitor engagement.

Table 4.3: Evaluation of ML Models for Predicting Pilgrimage Event Attendance

Model	Accuracy Score	Hamming Loss	Jaccard Score (Macro)	Samples F1 Score	
GRADIENT BOOSTING	0.495	0.250	0.654	0.77	Best overall: highest performance across metrics
RANDOM FOREST	0.469	0.291	0.604	0.70	Solid all-rounder, good baseline
K-NN	0.431	0.297	0.597	0.71	Consistent, good for nearby patterns
DECISION TREE	0.446	0.294	0.593	0.68	Simple & interpretable, but slightly lower scores
LOGISTIC REGRESSION	0.360	0.291	0.604	0.70	Fast, but misses some predictions entirely
SVM	0.351	0.321	0.580	0.72	High recall for some labels, but lower overall match

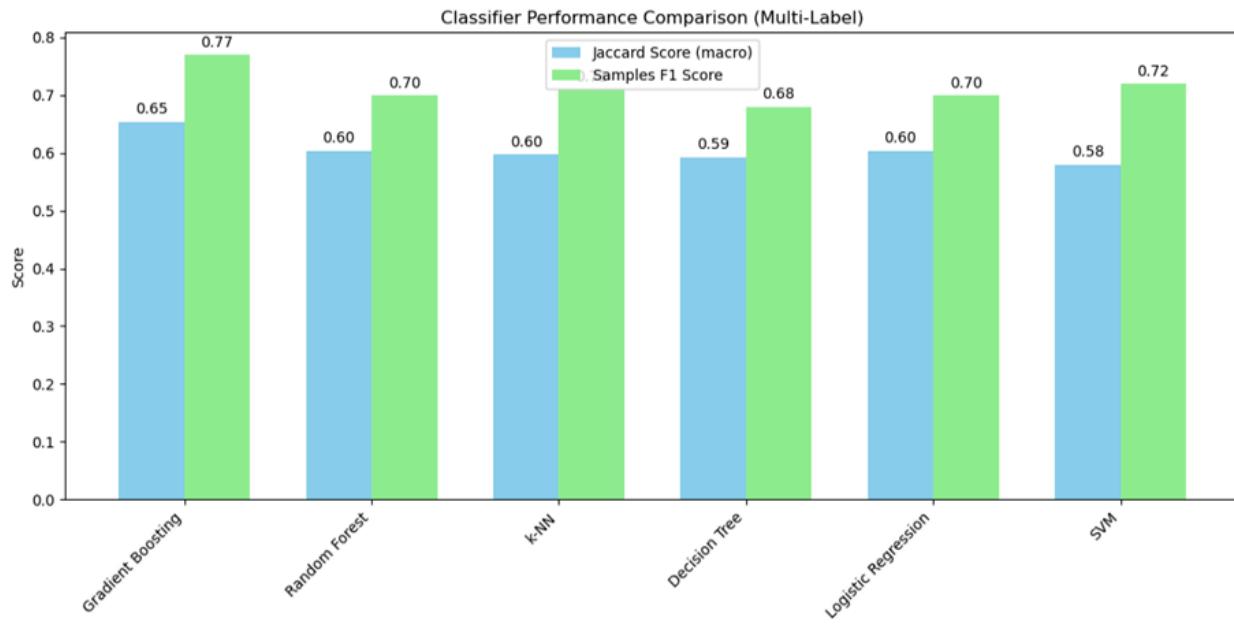


Fig.46. Comparison of Model performance

Random Forest and KNN are strong alternatives, offering reliable performance but with some trade-offs in speed and interpretability. Decision Tree provides simplicity and interpretability but at the cost of potential overfitting. Logistic Regression and SVM are less effective overall but might be worth considering for specific use cases where speed or recall for classes is essential.

3.7 Model Prediction

To make the machine learning model accessible to both researchers and practitioners, a user-friendly web interface was developed using Streamlit, a Python-based interactive application framework. The trained gradient boosting model, along with the label encoders for categorical variables, was serialized and loaded within the application. Users are prompted to enter relevant information including demographic attributes (e.g., age group, gender, state/country), visitation characteristics (e.g., type of visitor, camping status), and contextual factors (e.g., whether the visit falls on a weekend or during the Feast Day). Upon submission, the inputs are preprocessed using the corresponding label encoders and passed to the prediction model. The model outputs a multi-label prediction identifying the event(s) a visitor is most likely to attend. These events include Mass, Veneration of Relics, and Fair/Cultural Event/Sightseeing, and the predicted results are presented to the user in a readable format. This interactive tool demonstrates how machine learning can be operationalized to support visitor experience planning and religious tourism management at pilgrimage sites such as the Basilica of Bom Jesus.

Welcome to the First-Time Visitor Event Predictor!
Please enter values from the valid options shown below.

Age groups options: ['early professionals', 'mature adults', 'senior citizen', 'young adults', 'young dependents', 'youth']
 Gender: options: ['Male', 'female']
 State/Country options: ['andhra pradesh', 'arunachal pradesh', 'assam', 'australia', 'belarus', 'bihar', 'chhattisgarh', 'delhi', 'fiji', 'france', 'germany', 'goa', 'gujarat', 'haryana', 'himachal pradesh', 'ireland', 'jammu & kashmir', 'jharkhand', 'karnataka', 'kerala', 'khazakhstan', 'madhya pradesh', 'maharashtra', 'manipur', 'meghalaya', 'nagaland', 'nepal', 'netherlands', 'odisha', 'philippines', 'poland', 'portugal', 'punjab', 'rajasthan', 'russia', 'singapore', 'spain', 'switzerland', 'tamil nadu', 'tanzania', 'telangana', 'uk', 'usa', 'uttar pradesh', 'uttarakhand', 'washington dc', 'west bengal', 'zimbabwe']
 Type of visitors options: ['first time visitors', 'frequent visitors', 'long-time devotees', 'occasional visitors']
 Are you camping at the Basilica? options: ['no', 'yes']
X Invalid input. Please enter exactly one of the shown options.

Prediction complete!

* Predicted Events :
 Fair/Cultural Event/Sightseeing
 Veneration of Relics
 Mass

Fig.47. Prediction Interface using selected model

Basilica Event Guide

Select Age Group

early professionals

Select Gender

Male

Select State/Country

ireland

Type of Visitor

first time visitors

Are you camping at the Basilica?

no

Is it a weekend?

Yes

Is it the Feast Day?

Yes

Predict Events

Predicted Event(s): Fair/Cultural Event/Sightseeing

Fig.48. Interactive User Interface for Event Prediction

5. Discussions

Influence of Demographic Factors on Pilgrimage Participation (RQ1 / H1)

This study confirms that demographic factors such as age, gender, and region significantly influence participation in the Exposition at the Basilica of Bom Jesus. Chi-square tests revealed a strong association between age groups and event attendance, with older pilgrims more engaged in religious rituals such as Mass and relic veneration. Younger attendees, in contrast, tended to prefer cultural tourism activities. Gender analysis showed a slight inclination of females toward religious events, while males exhibited more interest in cultural programs. Visitors from different regions, including various states of India and overseas, also demonstrated varied participation patterns, shaped by cultural and geographic contexts. These findings support the hypothesis (H1) that demographic variables significantly influence pilgrimage engagement.

Engagement Patterns of First-Time vs. Repeat Visitors (RQ2 / H2)

The hypothesis that first-time visitors show comparatively different behavioral patterns to long-time devotees and frequent pilgrims is well supported by the data. It was observed that First-time visitors showed better engagement in sightseeing and cultural events, suggesting them being more touristy. Repeat visitors, on the other hand, demonstrated deeper spiritual engagement, with higher participation in Mass and veneration rituals. This distinction in participation highlights the need for differentiated planning in religious tourism, with targeted messaging and experiences based on visitor types.

Transformations in Pilgrimage Patterns Post-COVID (RQ3)

The COVID-19 pandemic significantly changed pilgrimage patterns. During the pandemic years, especially 2020 - 2021, mass gatherings had been stopped, and hence many pilgrims turned to online platforms for devotional engagement. YouTube data (Fig II.1) indicates a sharp spike in virtual participation, with views peaking at over 138,000 in 2021. However, the 18th Exposition in 2024 marked a strong return to in-person pilgrimage, drawing over 7 million attendees (Fig II.5). This revival reflects not only a return to tradition but also the persistence of hybrid participation. The pandemic has thus expanded pilgrimage modes, leading to a lasting shift toward dual digital-physical religious experiences.

The Role of YouTube and Digital Platforms (RQ4)

Digital platforms, especially YouTube, have emerged as key tools in shaping travel decisions and pilgrim behavior. The official Basilica channel recorded over 212,000 views for the 2024 opening ceremony, indicating strong global interest. Peaks in viewership aligned with major ceremonies and multilingual masses, demonstrating the platform's inclusivity and accessibility. First-time visitor engagement also rose significantly, likely influenced by digital exposure to the event. The data suggest that YouTube not only offers an alternative to physical attendance but

also serves as a powerful promotional and devotional medium, especially for diaspora Catholics and culturally curious audiences.

The 2024 Exposition's Impact on Goa's Tourism (RQ5 / H3)

Contrary to the hypothesis (H3) that the Exposition had no significant effect on Goa's tourist arrivals, 2024 recorded over 10.4 million visitors, the highest in the state's history (Appendix II). A significant surge was observed during the October–December quarter, coinciding with the Exposition period. The turnout of over 7 million pilgrims (TOI, 2024) underscores the Exposition's major contribution to tourism revival in Goa post-COVID. These findings refute the null hypothesis and affirm the Exposition's strong correlation with the 2024 tourism boom.

Predictive Modeling of Pilgrimage Participation (RQ6)

Supervised machine learning models were developed to predict event participation based on survey data. Gradient Boosting emerged as the most effective classifier. Variables such as "Type of Visitor," "State/Country," and "Camping at Basilica" were strong predictors, as reflected in high Cramér's V values. These models demonstrated high accuracy and interpretability, supporting their application in real-world religious tourism management, from anticipating crowd flows to customizing pilgrim experiences.

Key Insights

This analysis highlights how pilgrimage behavior at the Basilica of Bom Jesus has evolved into a multifaceted experience shaped by tradition, digital innovation, and demographic shifts. While sacred rituals like veneration of relics remain central, emerging trends such as increased online engagement, and culturally motivated travel reflect a broader redefinition of pilgrimage.

Accommodation trends also reflect this diversity. Budget-friendly barracks and camps catered to walking pilgrims, many of whom preserved their ancestral practices. At the same time, digital access expanded the Basilica's reach to those unable to attend in person, creating a hybrid devotional model.

Overall, this study illustrates how data-driven insights can inform religious event planning and deepen our understanding of how modern pilgrimages honor tradition while embracing technological change.

6. Conclusion

This research presents a data-driven exploration of pilgrimage patterns during the 18th Exposition of the Sacred Relics of St. Francis Xavier at the Basilica of Bom Jesus, Goa. Through survey data, viewership metrics, and machine learning models, the study identifies key trends influencing religious tourism in a post-pandemic world.

Findings show that pilgrimage today is no longer confined to physical presence it includes immense digital participation, cross-generational engagement, and culturally diverse motivations. By integrating quantitative techniques with cultural understanding, this research provides tourism authorities, religious institutions, and policymakers with actionable insights for future planning.

The Basilica emerges as a prime example of how heritage sites can remain relevant by embracing innovation while preserving their sanctity. Its dual identity as a religious and cultural landmark allows it to foster spiritual devotion and global connectivity.

Looking ahead, the deployment of predictive models and visitor segmentation tools can empower event organizers to design more inclusive, responsive, and sustainable pilgrimage experiences. Efforts should also be made to extend outreach to underrepresented regions and age groups and enhance infrastructure for a seamless pilgrim experience.

This study stands as a testament to the evolving narrative of pilgrimage rooted in tradition, informed by data, and inspired by the enduring human quest for connection, meaning, and faith.

7. References

- Carvache-Franco, M., Regalado-Pezúa, O., Carvache-Franco, O., & Carvache-Franco, W. (2024). Segmentation by motivations in religious tourism: A study of the Christ of Miracles Pilgrimage, Peru. *Plos one*, 19(5), e0303762.
- Kizhakanveatil Bhaskaran Pillai, S., Kamat, K., Scaglione, M., Pillai, B., & Kizhakanveatil, S. (2017). Exposition of St. Francis Xavier's Holy Relics in Goa: An Importance-Performance Analysis. *International Journal of Religious Tourism and Pilgrimage*, 5(1), 7. <https://doi.org/10.21427/D7TD8T>
- Dowson, R. (2020). A Discussion of the Practical and Theological Impacts of COVID-19 on Religious Worship, Events and Pilgrimage, from a Christian Perspective. *International Journal of Religious Tourism and Pilgrimage*, 8(7), 5. <https://doi.org/10.21427/ytdq-3s12>
- Marine-Roig, E. (2015). Religious Tourism versus Secular Pilgrimage: The Basilica of La Sagrada Família. *International Journal of Religious Tourism and Pilgrimage*, 3(1). <https://doi.org/10.21427/D7KM6Z>
- Raj, R., & Griffin, K. A. (2020). Reflecting on the Impact of COVID-19 on Religious Tourism and Pilgrimage. *International Journal of Religious Tourism and Pilgrimage*, 8(7) <https://doi.org/10.21427/8f91-6z16>
- Abiola-Oke, E., & Osioye, T. (2024). Hospitality Services and the Promotion of Religious Tourism: A Study of Redemption City, Nigeria. *International Journal of Religious Tourism*

Stamenkovic, Igor; Dodig, Ivan; and Vukojević, Dajana (2024) "Impact Analysis of Religious Events: Medjugorje as a Pilgrimage Destination," International Journal of Religious Tourism and Pilgrimage: Vol. 12: Iss. 2, Article 6

Collins-Kreiner, N. (2010) Researching Pilgrimage. Continuity and transformations. Annals of Tourism Research 37(2): 440–456

Agarwal, A., Kapoor, K., & Walia, S. (2021). Impact of Social Media on Spiritual Tourism in India: An SEM Analysis of the Critical Factors Impacting on Decision Making. International Journal of Religious Tourism and Pilgrimage, 9(5), 96–107. <https://doi.org/10.21427/1SCQ-MC52>

Bond, N., Packer, J., & Ballantyne, R. (2015). Exploring Visitor Experiences, Activities, and Benefits at Three Religious Tourism Sites. International Journal of Tourism Research, 17(5), 471–481. <https://doi.org/10.1002/jtr.2014>

Gokhale, K., Sawant, N., & Ugavekar, N. (2014). Goa: Tourism and Socio-Cultural Implications A Study Of Selected Tourist Destination of South Goa District, Goa-India. In IOSR Journal Of Humanities And Social Science (IOSR-JHSS) (Vol. 19, Issue 10). Ver. I. www.iosrjournals.org

Henn, A. (2008). Crossroads of religions: Shrines, mobility and urban space in Goa. International Journal of Urban and Regional Research, 32(3), 658–670. <https://doi.org/10.1111/j.1468-2427.2008.00803.x>

Mukherjee, S., Bhattacharjee, S., & Singha, S. S. (n.d.). Religious to Spiritual Tourism- An Era of Paradigm Shift in India. <https://ssrn.com/abstract=3546903>

Tan, J. (2018). Religious tourism and the new evangelization: theory and best practice in the pastoral promotion of the Church's cultural heritage. Church, Communication and Culture, 3(3), 283–309. <https://doi.org/10.1080/23753234.2018.1547117>

Nolan, M. L., and S. Nolan. (1992). "Religious Sites as Tourism Attractions." Annals of Tourism Research 19 (1): 68–78. doi:10.1016/0160-7383(92)90107-Z

Irimias, Anna; Mitev, Ariel; and Michalko, Gabor (2016) "Demographic Characteristics Influencing Religious Tourism Behaviour: Evidence from a Central-Eastern-European country," International Journal of Religious Tourism and Pilgrimage: Vol. 4: Iss. 4, Article 3. doi:<https://doi.org/10.21427/D7VB0D>

Dein, S., & Watts, F. (2023). Religious worship online: A qualitative study of two Sunday virtual services. Archive for the Psychology of Religion, 45(2), 191-209. <https://doi.org/10.1177/00846724221145348>

de Azevedo, C. (1956). The Churches of Goa. In Source: Journal of the Society of Architectural Historians (Vol. 15, Issue 3)

Appendix I

Survey Form

The following questions were used in the survey conducted with pilgrims visiting the Basilica of Bom Jesus.

- Age:

- Gender:

Male

Female

Others

- State of Residence:

- Event Attending:

Veneration of Relics

Mass

Fair/ Cultural Event/ Sightseeing

- Are you staying over(camping) at the Basilica?

Yes

No

- Since how many years have you been visiting the Basilica?

Its my first time

Every year since birth

5 years

Other _____

- I consent to my information being used for research purposes only

Yes

No

Survey for the St. Francis Xavier Exposition 2024					
Age	Gender	Indian State of Residence/ Country	Event Attending?	How many years have you been visiting the Basilica? (approx)	I consent to my information being used for research purposes only.
	<input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Other		<input type="checkbox"/> Veneration of Relics <input type="checkbox"/> Mass <input type="checkbox"/> Fair/ Cultural Event / Sightseeing	<input type="checkbox"/> Its my first time <input type="checkbox"/> _____ years	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Other		<input type="checkbox"/> Veneration of Relics <input type="checkbox"/> Mass <input type="checkbox"/> Fair/ Cultural Event / Sightseeing	<input type="checkbox"/> Its my first time <input type="checkbox"/> _____ years	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Other		<input type="checkbox"/> Veneration of Relics <input type="checkbox"/> Mass <input type="checkbox"/> Fair/ Cultural Event / Sightseeing	<input type="checkbox"/> Its my first time <input type="checkbox"/> _____ years	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Other		<input type="checkbox"/> Veneration of Relics <input type="checkbox"/> Mass <input type="checkbox"/> Fair/ Cultural Event / Sightseeing	<input type="checkbox"/> Its my first time <input type="checkbox"/> _____ years	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Other		<input type="checkbox"/> Veneration of Relics <input type="checkbox"/> Mass <input type="checkbox"/> Fair/ Cultural Event / Sightseeing	<input type="checkbox"/> Its my first time <input type="checkbox"/> _____ years	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Other		<input type="checkbox"/> Veneration of Relics <input type="checkbox"/> Mass <input type="checkbox"/> Fair/ Cultural Event / Sightseeing	<input type="checkbox"/> Its my first time <input type="checkbox"/> _____ years	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Other		<input type="checkbox"/> Veneration of Relics <input type="checkbox"/> Mass <input type="checkbox"/> Fair/ Cultural Event / Sightseeing	<input type="checkbox"/> Its my first time <input type="checkbox"/> _____ years	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Other		<input type="checkbox"/> Veneration of Relics <input type="checkbox"/> Mass <input type="checkbox"/> Fair/ Cultural Event / Sightseeing	<input type="checkbox"/> Its my first time <input type="checkbox"/> _____ years	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Other		<input type="checkbox"/> Veneration of Relics <input type="checkbox"/> Mass <input type="checkbox"/> Fair/ Cultural Event / Sightseeing	<input type="checkbox"/> Its my first time <input type="checkbox"/> _____ years	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Yes <input type="checkbox"/> No

Fig. I.1 : Offline survey sheet that was used to collect data from visitors at the Basilica



Research Survey for the St. Francis Xavier Exposition 2024

msds.2314@unigoa.ac.in [Switch account](#) 

 Not shared

* Indicates required question

Your Age *

Your answer

Gender: *

Female

Fig. I.2: Google form that was used to collect data from visitors through online mode

Appendix II

From Decline to Revival: The Role of the 18th Exposition of relics of St. Francis Xavier in Goa's Tourism Revival Post-COVID-19

Although this dataset was not directly used in the core analysis of this study, it has been included here to provide additional context regarding broader tourism trends in Goa during the years surrounding the Exposition.

The global outbreak of COVID-19 in early 2020 had a profound impact on international and domestic travel. Goa, one of India's most popular tourist destinations, witnessed a sharp decline in tourist arrivals, particularly by flight and cruise.

However, 2024 has witnessed a remarkable recovery in tourism, with over 10.4 million visitors recorded surpassing all previous full-year records and marking the highest tourism influx in Goa's recent history. This reflects not just a return to normal but a renewed surge in travel interest, influenced by resumed religious events and shifting post-pandemic travel patterns.

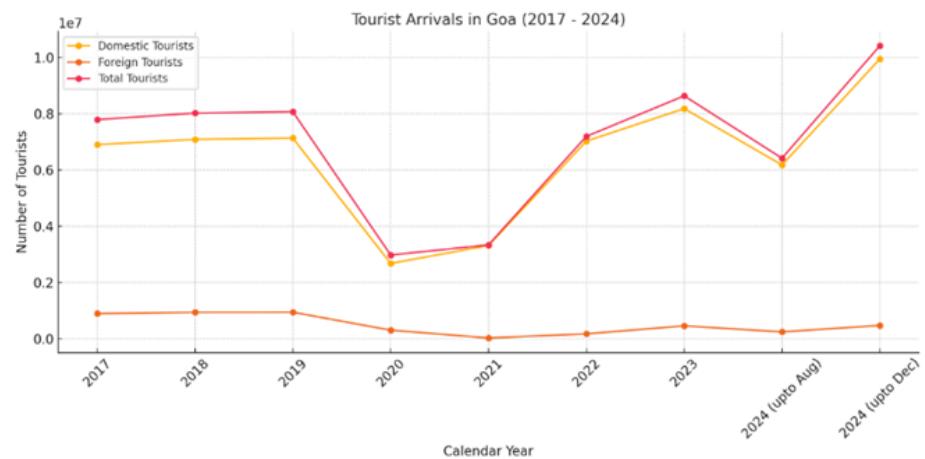


Fig. II.1: Tourist arrivals in Goa from 2017-2024 including local and foreign tourists

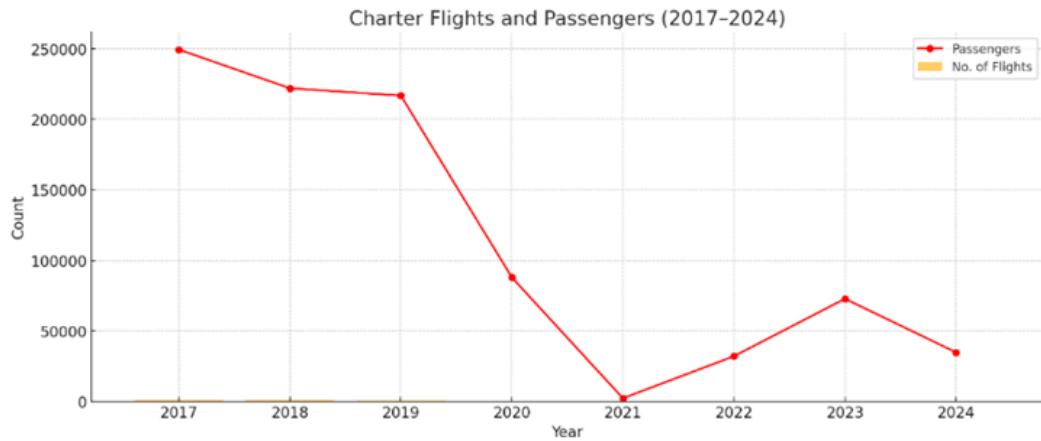


Fig. II.2: Charter flights in Goa from 2017-August 2024 including local and foreign tourists

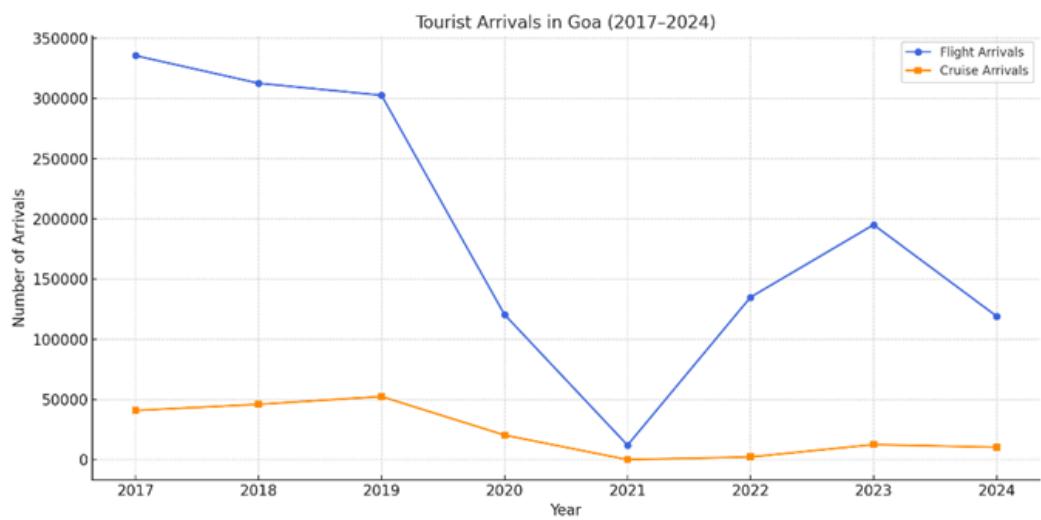


Fig. II.3: Flight and Cruise Arrivals in Goa (2017–August 2024)

From 2017 to 2019, Goa experienced a stable inflow of tourists arriving via both flights and cruises (Fig II.3), with an average of over 300,000 passengers by air and 45,000 cruise tourists annually. These figures reflected Goa's strong appeal as a leisure, cultural, and religious tourism hub.

The COVID-19 pandemic disrupted this trajectory. Flight arrivals dropped by more than half, and cruise tourism completely ceased in 2021. Only 11,971 passengers arrived by flight that year, and cruise arrivals were zero. These declines mirror global travel disruptions due to health protocols, lockdowns, and port closures.

A gradual recovery began in 2022 and picked up momentum in 2023 as international charters resumed and domestic tourism rebounded. While total flight and cruise traffic by August 2024

still lagged pre-pandemic benchmarks, a sudden surge was recorded in the final quarter of the year, particularly in December 2024, aligning with major religious festivities and the peak Exposition period.

The Times of India (The Times of India, 2025), reported that the Oct-Dec quarter saw a growth of 38%, with domestic tourists increasing by 22% and foreign tourists by 3% in 2024.

EDITION IN MUMBAI 31°C

THE TIMES OF INDIA SUBSCRIBE

City Goa Mumbai Delhi Bengaluru Hyderabad Kolkata Chennai Agra Agartala Ahmedabad Ajmer Amaravati

GOA ELECTIONS PHOTOS

NEWS / CITY NEWS / GOA NEWS / In Contrast To Claims, 21% Growth In Tourist Arrivals In '24: G...

TRENDING Delhi Assembly Election Winner List Maha Kumbh Chhattisgarh Unnatural Sex Law Karnataka Murder Ti < >

In contrast to claims, 21% growth in tourist arrivals in '24: Goa govt

Times News Network / TNN / Jan 14, 2025, 00:26 IST

SHARE AA FOLLOW US



Panaji: Goa recorded a robust 21% jump in tourist arrivals in 2024, largely driven by domestic tourists who flocked to the state in the Oct-Dec quarter. The state, whose population is estimated at 15.7 lakh, saw over 1 crore tourists enter the state in the last 12 months. In 2023, the total tourist arrivals were pegged at 86.3 lakh.

According to provisional data released by the tourism department on Monday, Goa received 99.4

Fig. II.4: Oct-Dec quarter saw a growth of 38% in tourist arrivals

At 70 lakh, Old Goa sees record pilgrim turnout for Exposition

TNN / Dec 31, 2024, 23:46 IST



TOI

Panaji: At least 70 lakh pilgrims visited Se Cathedral to view and venerate the relics of St Francis Xavier. With the crowds seen during the last few days, the numbers are expected to cross 80 lakh, deputy chairman of the Exposition committee, minister Aleixo Sequeira, said.

"It's not that we set a target. People are coming on their own. Nobody forced them to come to Old Goa. They are doing it out of respect and fervour

for St Francis Xavier. These are not only people from Goa but also from various parts of the world," he said, adding that many decided to venerate the relics during the fag end of the Exposition.

The highest footfalls were seen on Dec 27 and 28, with people spilling over onto the road in the heat of the sun. "It was a pity to see people standing for hours," convener of the Exposition committee, Fr Henry Falcao, said. "The crowds are still there, but they are manageable. Right from 6am to 7pm, there is a continuous flow of people."

Liaison officer between govt and the Exposition committee, Fr Lawrence Fernandes, has appealed to the authorities to preserve Old Goa and its sanctity and hygiene even after the Exposition. "We must maintain this as a Unesco World Heritage Site," he said.

Fig. II.5: Over 70 Lakh Pilgrims Visit Old Goa During the Exposition

According to The Times of India (The Times of India, 2024), the 18th Exposition of the Relics of St. Francis Xavier was a major religious event that drew at least 7 million pilgrims to the Se Cathedral in Old Goa, with expectations that the number would surpass 8 million (80 lakh) by the event's conclusion.

While TOI primarily highlights the scale of the Exposition itself, this massive turnout is likely to have been a significant contributor to the broader surge in tourism witnessed in Goa during 2024. The Exposition's enduring global appeal and the scale of participation underscore the growing role of religious tourism in shaping post-pandemic travel trends in the region.