Case Study Report



**Tech Saksham**

Data Analytics with Power BI

**“Real-Time Analysis of Bank Customers”**

**“ST. JOHN’S COLLEGE, PALAYAMKOTTAI”**

|  |  |
| --- | --- |
| **NM ID** | **NAME** |
| FE3FB39BB585ED6B8244369D43815A74 | MANIMEGALA. G |

|  |  |
| --- | --- |
|  |  |
|  | Trainer Name |
|  | Master Trainer |

**ABSTRACT**

In the digital age, data has become an invaluable asset for global terrorism dataset analysis, particularly in the global terrorism. The proposed project, “Power BI Powered Global Terrorism Dataset Analysis,” aims to leverage Power BI, a global terrorism dataset includes all events that meet its criteria for terrorism and are documented by at least one high-quality source. Allows filtering out any events that may not meet one’s own definition of terrorism by sharing disaggregated data on specific characteristics. The Global Terrorism Dataset (GDT) documents more than 200,000 international and domestic terrorist attacks that occurred worldwide since 1970. The Global Terrorism Index (GTI) report is produced by the Institute for Economics & Peace (IEP) using data from Terrorism Tracker and other sources. GTI, the Institute for Economics & Peace (IEP) publishes the Global Terrorism Index, which was created by Steve Killelea, an IT entrepreneur, and IEP founder.

**INDEX**

|  |  |  |
| --- | --- | --- |
| **Sr. No.** | **Table of Contents** | **Page No.** |
| 1 | Chapter 1: Introduction | 4 |
| 2 | Chapter 2: Services and Tools Required | 6 |
| 3 | Chapter 3: Project Architecture | 7 |
| 4 | Chapter 4: Modeling and Result | 9 |
| 5 | Conclusion | 18 |
| 6 | Future Scope | 19 |
| 7 | References | 20 |
| 8 | Links | 21 |

**CHAPTER 1**

**INTRODUCTION**

* 1. **Problem Statement**

The Global Terrorism Database (GTD) documents more than 200,000 international and domestic terrorist attacks that occurred worldwide since 1970. With details on various dimensions of each attack, the GTD familiarizes analysts, policymakers, scholars, and journalists with patterns of terrorism. According to Grob-Fitzgerald (2005), terrorism can be broken down into four broad forms: national terrorism – terrorist activities involving national borders; revolutionary terrorism – activities aimed at the philosophical and political nature of government; reactionary terrorism – activities concerned with preventing ...

* 1. **Proposed Solution**

In the solution of this problem, Power BI can be the terrorism, Systematic use of violence to create a general climate of fear in a population and thereby to bring about a particular political objective. It has been used throughout history by political organizations of both the left and the right, by nationalist and ethnic groups, and by revolutionaries. There are five crucial components of terrorism, an involvement of an act of violence, an audience, the creation of a mood of fear, innocent victims, and political goals or motive. The Global Terrorism Database--or GTD--began in 2001 when researchers at the University of Maryland obtained a large database originally collected by the Pinkerton Global Intelligence Services (PGIS).

* 1. **Feature**

**1.** The Global Terrorism Database--or GTD--began in 2001 when researchers at the University of Maryland obtained a large database originally collected by the Pinkerton Global Intelligence Services (PGIS)..

**2.** The National Terrorism Advisory System (NTAS) is designed to communicate information about terrorist threats by providing timely, detailed information to the American public.

**1.4 Advantages**

* Currently the most comprehensive unclassified database on terrorist attacks in the world
* Includes information on more than 88,000 bombings, 19,000 assassinations, and 11,000 kidnappings since 1970
* Includes information on at least 45 variables for each case, with more recent incidents including information on more than 120 variables
* More than 4,000,000 news articles and 25,000 news sources were reviewed to collect incident data from 1998 to 2017 alone
  1. **Scope**

The scope of this project extends to all banking institutions that aim to leverage data for decision-making and customer engagement. The project can be further extended to incorporate more data sources and advanced analytics techniques, such as machine learning and artificial intelligence, to provide more sophisticated insights into customer behavior. Terrorism is the use of violence and intimidation, especially against civilians, in the pursuit of political, ideological, or religious goals. It is a tactic used by individuals or groups to achieve their objectives by creating fear and causing disruption.

Terrorism can take many forms, including bombings, assassinations, hijackings, and cyberattacks. It can be carried out by state actors or by non-state actors, such as terrorist organizations or extremist groups. The impact of terrorism is far-reaching, as it can cause physical harm, psychological trauma, and economic damage.

However, the problem of terrorism remains a significant global challenge, and efforts to address it must be ongoing and multifaceted.

.

**CHAPTER 2**

**SERVICES AND TOOLS REQUIRED**

**2.1 Services Used**

* **Data Collection and Storage Services:** GTD includes all events that meet its criteria for terrorism and are documented by at least one high-quality source.
* **Data Processing Services**: With details on various dimensions of each attack, the GTD familiarizes analysts, policymakers, scholars, and journalists with patterns of terrorism.
* **Machine Learning Services**: Azure Machine Learning or AWS Sage Maker can be used to build predictive models based on historical data.

**2.2 Tools and Software used**

**Tools**:

* **Power BI**: The main tool for this project is Power BI, which will be used to create interactive dashboards for real-time data visualization.
* **Power Query**: This is a data connection technology that enables you to discover, connect, combine, and refine data across a wide variety of sources.

**Software Requirements**:

* **Power BI Desktop**: This is a Windows application that you can use to create reports and publish them to Power BI.
* **Power BI Service**: This is an online SaaS (Software as a Service) service that you use to publish reports, create new dashboards, and share insights.
* **Power BI Mobile**: This is a mobile application that you can use to access your reports and dashboards on the go.

**CHAPTER 3**

**PROJECT ARCHITECTURE**

**3.1 Architecture**

**USER FRONTEND BACKEND**

|  |  |  |
| --- | --- | --- |
|  | **HTML 5** | **NODEJS 14.0**  **Database** |

Here’s a high-level architecture for the project:

1. **Data Collection**: Real-time customer data is collected from various sources like bank transactions, customer interactions, etc. This could be achieved using services like Azure Event Hubs or AWS Kinesis.
2. **Data Storage**: The collected data is stored in a database for processing. Azure SQL Database or AWS RDS can be used for this purpose.
3. **Data Processing**: The stored data is processed in real-time using services like Azure Stream Analytics or AWS Kinesis Data Analytics.
4. **Machine Learning**: Predictive models are built based on processed data using Azure Machine Learning or AWS Sage Maker. These models can help in predicting customer behavior, detecting fraud, etc.
5. **Data Visualization**: The processed data and the results from the predictive models are visualized in real-time using Power BI. Power BI allows you to create interactive dashboards that can provide valuable insights into the data.
6. **Data Access**: The dashboards created in Power BI can be accessed through Power BI Desktop, Power BI Service (online), and Power BI Mobile.

This architecture provides a comprehensive solution for real-time analysis of bank customers. However, it’s important to note that the specific architecture may vary depending on the bank’s existing infrastructure, specific requirements, and budget. It’s also important to ensure that all tools and services comply with relevant data privacy and security regulations.

**CHAPTER 4**

**MODELING AND RESULT**

**Manage relationship**

The “disp” file will be used as the main connector as it contains most key identifier (account id, client id and disp id) which can be use to relates the 8 data files together. The “district” file is use to link the client profile geographically with “district id”







**Modelling for Gender and Age data**

Notice that the Gender and age of the client are missing from the data. These can be formulated from the birth number YYMMDD where at months (the 3rd and 4th digits) greater than 50 means that client is a Female. We can create a column for Gender.



For birthday, we need to reduce the birth month of the female by 50 and then change the date format to DD/MM/YYYY adding 1900 to the year.



For Age, we shall assume it is year 1999 as explain previously and use it to minus from the birth year.



**Replacing values**

Set some fields to English for easy understanding, we replace values to English with the Power Query Editor.







Changing the order of Region name at Power Query

Duplicate the “district /region” then split column using space as delimiter.



Then merge column by Region and direction. Refer to applied steps for details.



**Grouping of age by ranges**

As the customers’ age ranges from 12 to 88, we shall group them into different generation age range for easier profiling, we will group the ages into 5 groups.

The Gen Y are youths,

Gen X are young working adults, some starting their families

Baby Boomer are working adults with families.

The silent Generations some are working and retired, living on pensions.

The greatest Generation, retired elderly living on pensions.



**Credit Rating and Loan Status**

As the Loan status uses A, B, C, D which are not reader friendly. We can add a column to represent what it stands for, we also simplify the classification of those with late or default on payment as bad credit, refer to the table below for details on the new columns added.



Values of such as “account Id” have also been set as Text.

And District name have been categorized as place to be use for the map to show the sum of the inhabitants in each region.

**Dashboard**

A screenshot of a credit card

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a credit card

Description automatically generated

**CONCLUSION**

The project “Real-Time Analysis of Bank Customers” using Power BI has successfully demonstrated the potential of data analytics in the banking sector. The real-time analysis of customer data has provided valuable insights into customer behavior, preferences, and trends, thereby facilitating informed decision-making. The interactive dashboards and reports have offered a comprehensive view of customer data, enabling the identification of patterns and correlations. This has not only improved the efficiency of data analysis but also enhanced the bank’s ability to provide personalized services to its customers. The project has also highlighted the importance of data visualization in making complex data more understandable and accessible. The use of Power BI has made it possible to present data in a visually appealing and easy-to-understand format, thereby aiding in better decision-making.

**FUTURE SCOPE**

The future scope of this project is vast. With the advent of advanced analytics and machine learning, Power BI can be leveraged to predict future trends based on historical data. Integrating these predictive analytics into the project could enable the bank to anticipate customer needs and proactively offer solutions. Furthermore, Power BI’s capability to integrate with various data sources opens up the possibility of incorporating more diverse datasets for a more holistic view of customers. As data privacy and security become increasingly important, future iterations of this project should focus on implementing robust data governance strategies. This would ensure the secure handling of sensitive customer data while complying with data protection regulations. Additionally, the project could explore the integration of real-time data streams to provide even more timely and relevant insights. This could potentially transform the way banks interact with their customers, leading to improved customer satisfaction and loyalty.

**REFERENCES**

<https://medium.com/analytics-vidhya/analysis-of-bank-customers-using-dashboard-in-power-bi-a366f2b3e563>

**LINK**

<https://github.com/githubtraining/hellogitworld.git>