ITBN CONFERENCE – DATA ANALYSIS

- Adarsh BG
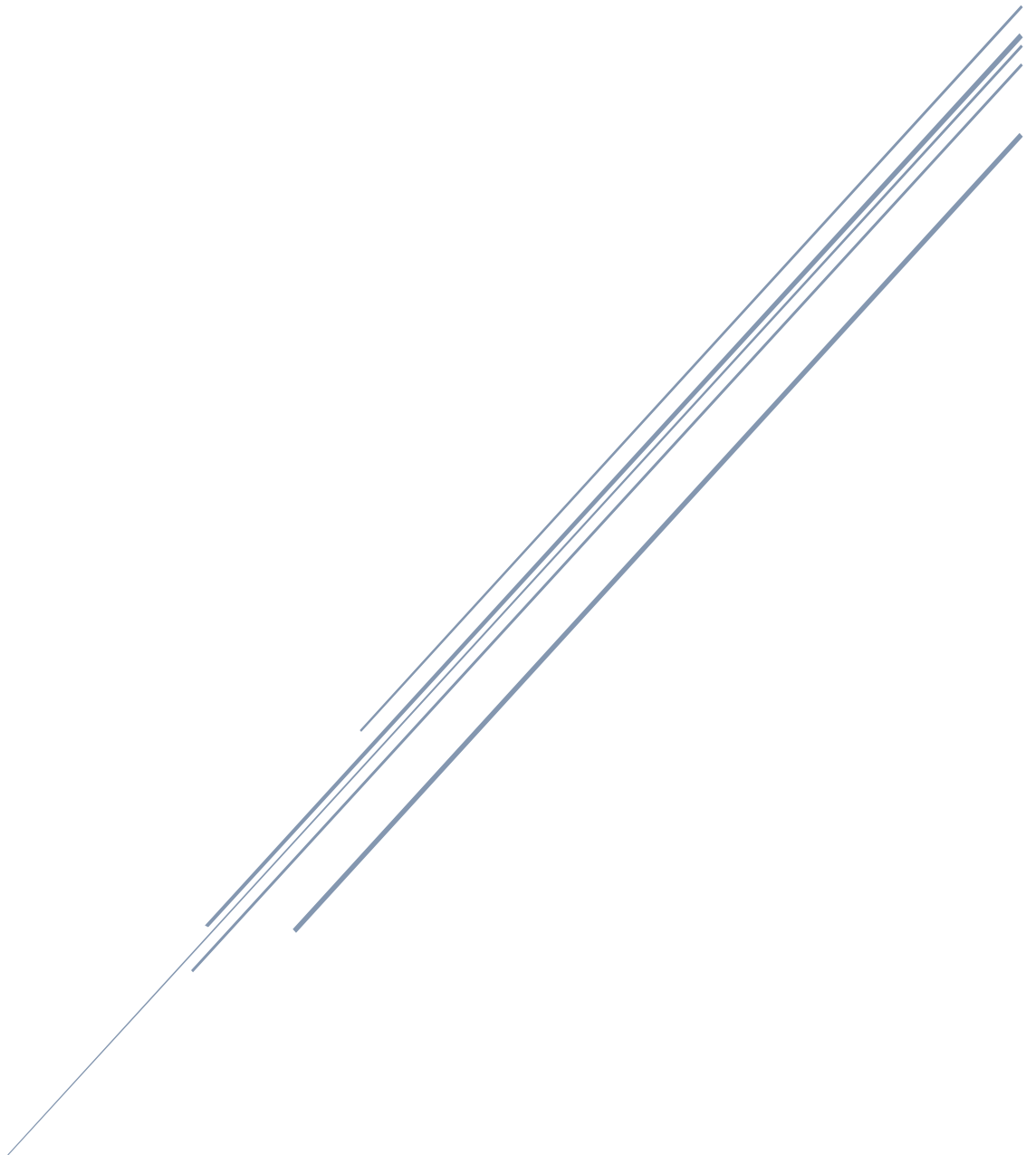
CEU
Business Analytics

Table of Contents

1. Introduction

ITBN is the biggest IT security conference in Central European region. Local representatives of the IT Security market collaborated in 2004 to create a forum to represent the interest of all the market players independently. The original goal has been far exceeded by ITBN: the event is the home for innovations and announcements of fresh products. ITBN presents the actual opinion of government officials and legal, financial, professional spheres of the IT Security profession.

ITBN has been carefully collecting the conference data for past few years and has stored in a database. Below are the details of datasets available:

1. **User Registration data**: Users details, personal information, professional background, company details etc.
2. **Presentation/Talks**: Presentations done on each conference year, topics, presenter, links to presentation video, documents etc.
3. **QR Badges data**: Each conference attendee was given RFID tag and their movement and presence in different presentations were tracked with this. Userid, SwipeIn/SwipeOut data was available.
4. **Survey Data** – This is the highlight of all the data collected. Each user was supposed to take up survey and give information related to:
   a. Their organization, size, employee count etc.
   b. IT Security satisfaction Index.
   c. IT security Budget, investment variance.
   d. Vendors used etc.

All this data was distributed in different tables in database and was never really analysed.

The idea was to analyse this data and bring out the underlying insights, trends relation between different factors etc.

Initial scope of Analysis:

1. Perform user profiling based on the information that they have provided and adding some external data and do user segmentation. Based on the user segment, their interest and background build recommendation engine which will suggest users to attend presentations which they might be interested in. Incorporate this recommendation engine in the ITBN conference app for 2018.
   > Due to data security constraints all user specific information was anonymized, so we could not perform this activity.
2. Analyse Survey data and understand the current IT security market trends in this region, how organizations are evolving, how professionals are equipping themselves with latest trends.
3. How IT security budget is varying in organizations, identifying potential customers who has reported high investments in IT security.
4. Analysing QR Badge data to understand what the most popular presentations are based on the number of users attended, determining appropriate timing for each presentation which will convenient for both Presenters and Attendees.

This analysis was presented in ITBN conference 2018 on Sep 26th at Groupama Arena, Budapest.

2. Analysis and Visualization:

The data was available in MySQL dump. As data was already structured in relational database tables, I used predominantly SQL for joining data in different tables and analysing. I have used Tableau and MSBI for further visualizations and Reporting.

All Tableau and MSBI reports are interactive, we can look for the details in specific Industry and look for the data reported by professionals of specific roles like 'IT Security' specialist etc.

Tableau report was published in Tableau Public online server, below is the link. (This report may not be available for many days as it is published through Tableau trial version).

https://eu-west-1a.online.tableau.com/#/site/adarshbg/views/ITBNDataAnalysis/OrganizationTrend?:iid=3

2.1 Organization Trend:

This dashboard shows how organizations are evolving, from what kind of organizations users are attending ITBN conference etc.
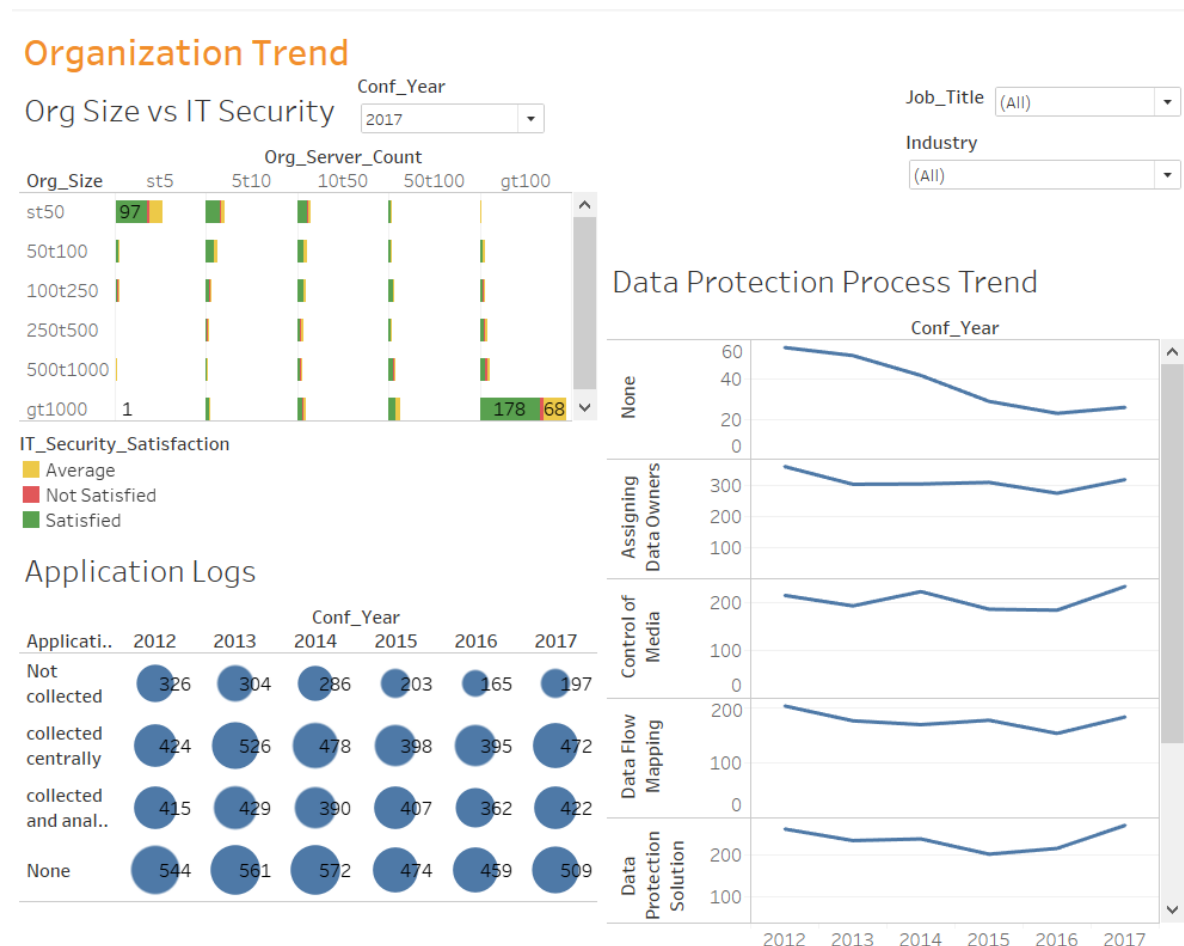

Figure 1: Organization Trend - Tableau Report

- **Org Size vs IT Security**: Most of the conference attendees are from either very small organizations or very big organizations. Looks like majority of them are satisfied with their company's IT security setup.

- **Application logs usage**: Collecting and analysing application logs is very important to proactively avoiding security incidents and to improve performance. Looks like not much progress is happening in this aspect, as industry is moving towards predictive maintenance and all, we can expect more effort here.
- **Data Protection Processes**: If we see yearly trend, people who are not following any data protection processes is considerably reducing. There is a steep increase in the companies following various DP processes between 2016 to 2017, it may be because of GDPR implementation.

2.2 Visitors trend:

Similar analysis was done conference attendees, below are some insights:

1. Majority of the people were having 'Szakember' (IT Security specialist).
2. Large number of attendees have CISA certification. People with CISSP certification has steeply increased from 2016 to 2017.
3. Attendees have shown more interest on Virus, Malware, Virtual cloud recently.

2.3 IT Security Budget:

Majority of the attendees reported no change in the IT security budget. People reporting small or major increase in budget is increasing over years. People reporting decrease in budget is reducing over years.

Further analysis was not possible as users have not entered the details in free text field consciously.

2.4 Vendor details:

Word cloud was prepared to identify most popular vendors. Again, this is interactive report which can be filtered for specific sector or industry.

2.5 MSBI Reports:

Two MSBI reports were prepared. Data is sliced and diced at various levels. People who have minimal experience working with Excel Pivot can easily drill up/drown through the report.

1. **Popular Presentations:** This shows how many users attended each presentation, how long they stayed etc. Winners or popular presentations are related to Virus Malware, Virtual cloud and so on.
2. **Year to Year Budget growth:** This report is to identify users who has reported high budget growth. Due to data quality issues this is not very accurate.

2.6 Miscellaneous analysis:

I tried to add in external data to explore more, below are some insights:

1. **Adding Weather data:** Added weather data to see the its impact on the conference attendance. As weather stayed pretty good on all year's conference dates could not notices any impact of it. Only day it was raining but there was no significant change in the conference attendance.
2. Tried to compare the attendees demand curve for the various IT security related topics with the **Gartner Hype cycle reports**. Could not see much correlations.

3. Next steps:

As the data was anonymized due to data security constraints, major part of analysis could not be done. Very few, around 100% of the attendees filled survey properly, this was another challenge for analysis.

ITBN team is now working on preparing GDPR ready solution for Data analysis. We are hoping to get the original data (without anonymization) and continue with the initial scope of analysis.