

A person wearing a blue protective suit, mask, and cap is walking away from the camera down a brightly lit hospital hallway. The floor is highly reflective, showing the person's reflection. The hallway has white walls and windows on the left. Several vibrant, out-of-focus light trails in shades of blue, green, and yellow streak across the scene, creating a sense of motion and technology.

# **JOINT COMMISSION ACCREDITATION SURVEY READINESS - EXECUTIVE SUMMARY**

**DIZ-E DOCTORS CONSULTING, LLC**

## EXECUTIVE OVERVIEW

For the past ten weeks, Diz-E Doctors Consulting has been analyzing hospital survey data from The Joint Commission (TJC). This executive summary will provide a high level overview of the completion of the project. We will outline our problem statement below along with a summary of our approach to exploratory data, data transformation and modeling. The final section of the summary includes our conclusions, recommendations, and next steps to improve the findings going forward. The deliverables include various visualizations, a predictive model to assist hospitals identify probable findings, a dashboard illustration, an interactive mobile application as well as informational materials for education and training of TJC employees.

## PROBLEM STATEMENT

For the project, we wanted to transform the TJC's scattered and unstructured data into a meaningful display of risks faced by each hospital in the United States. This will enable TJC to redesign their survey process to focus on the "high-risk, high-yield" areas during the survey process that will most likely impact patient safety. This valuable resource can also be shared with the healthcare institutions whose financial solvency depends on a successful TJC survey.

## ASSESSMENT

### **Data**

The data scrubbing and structuring process from the three separate TJC databases took a large amount of project resources. This will need to be repeated during each iteration of data extraction until TJC updates and standardizes their data.

Since the data for this analysis comes from three separate Joint Commission databases across three divisions (Survey Findings, Survey Event, Main Site/Secondary Site Characteristics), a significant opportunity for our company exists to provide TJC a real-time solution. We possess distinctive competencies with industry leading practices through the use of Amazon Web Services coupled with MySQL Workbench for data storage, manipulation and shared (secure) access. There was a seamless transfer between the data and analysis sections, and all tables were available real time to all team members. SAFER scores were put into practice in late 2016 and are not available on results data from prior years.

An opportunity exists in re-exploring additional SAFER survey data since our analysis reveals a significant change in the survey findings and processes over time. TJC leadership was also very excited about the demonstration of the mobile application saying *"This is DEFINITELY more on par with what I was thinking about when you started this project. I really like the concept of selecting a standard and then seeing what the probability of findings on other standards would*

be." "I agree...this is really cool!!! Customers would go crazy over this...it definitely shows the Power of Data."

Please take the time to review this significant product opportunity:

[https://predict498-teamfour.shinyapps.io/health\\_care\\_dashboard/](https://predict498-teamfour.shinyapps.io/health_care_dashboard/)

Read below for some more detail on this app.

### **Modeling**

One major area of focus was to predict the SAFER score that an organization was likely to incur based on hospital type and Standard findings. Recall that a severe finding could result in a SAFER score of 10 which could cause an organization to lose accreditation. The goal is to predict the range of SAFER scores that an organization is likely to receive based on previous surveys.

Several modeling methods were attempted, including

- Linear regression models
- Decision trees
- Gradient boosting models
- Random forests for regression
- Support vector machines [SVM]
- Generalized linear models [GLM]
- Random forests for multiclass classification

With 26 distinct levels for the hospital type variable, the performance and reliability of the SVM and GLM methods for multiclass classification were less than satisfactory.

The existence of clusters that are not differentiated by demographics but by findings suggest that something else is creating these outcomes. Possible factors include:

- geographic location
- individual surveyor(s) responsible for survey
- hospital staffing mix
- countless other possibilities

There is some concern regarding the dimensionality of the data. This is particularly true of the Standards and EPs. Each of the roughly 200 Standards belongs to one of 18 Standard Chapter Names, and models built on Standard Chapter Names do not have enough information to detect the importance of the Standards. Models built using the individual Standards treat each Standard as an individual entity however, and a middle level between Standard and Standard



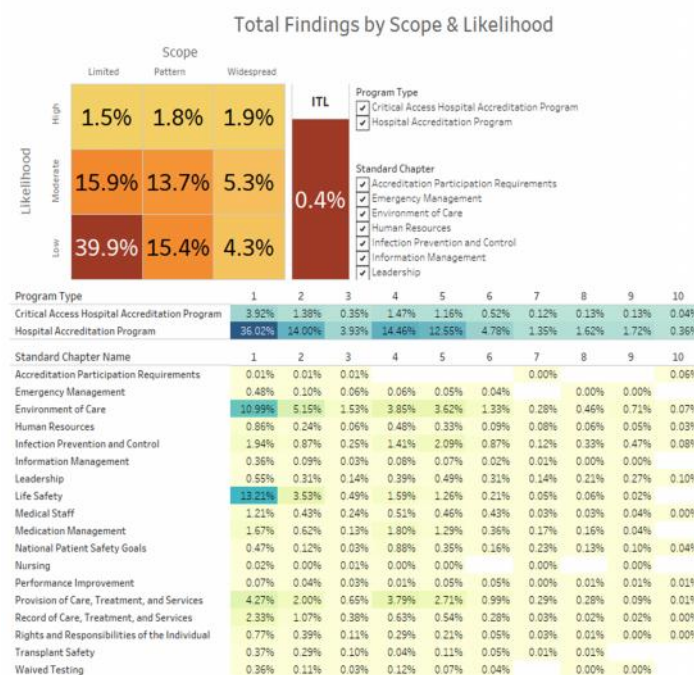
Chapter Name that groups Standards by their similarities may provide the right level of detail to use Standards as a useful and accurate predictor.

## PRODUCT DESCRIPTIONS

### **Dashboard**

A descriptive analytics visualization tool (a.k.a. dashboard) was used as an important tool for project data analysis. The dashboard combined electronic survey findings data with hospital demographic data including bed size, average daily census, number of care sites and visits to provide a better understanding of the historical patterns and correlations between survey findings, severity score and hospital demographics. The final dashboard will be provided as a tool for similar use by TJC decision makers and optionally the hospital systems in which surveys are performed. Leveraging such a visualization tool should enable quick analysis and effective presentation of trends to important stakeholders and clients of TJC.

The dashboard depicted below shows a survey SAFER score grid associated with the findings based on scope and likelihood of each finding's score. The scores range from 1 – 10, with 10 representing an Immediate Threat to Life (ITL) – the most severe score and could cause a loss of accreditation or complete shut down to occur. More information on TJC's SAFER Matrix scoring can be found at the TJC website. This dashboard is available for viewing [here](#).



## Mobile App

R Shiny and Amazon RDS were combined to develop an application that will be available to TJC staff and healthcare organizations to assist in the preparation of survey events.

The figure below provides a view of the Standards tab. From this tab, the user can select a hospital type and a Standard and see the minimum, average, and maximum SAFER Scores predicted for that hospital type and Standard pair. The plot shows the top five Standards and their probabilities that are likely for the Standard and hospital pair. From the screenshot below it is clear that for an Academic healthcare organization an Infection Prevention and Control - IC.02.02.01 is a reasonably serious finding and there is at least a 30% chance of incurring three of the five Standards listed in the plot. The application runs on desktops, tablets and cell phones and connects to the Amazon RDS instance for data retrieval.

