Judging Guidelines

Best Insight

- All teams that present their work will be eligible for the *Best Insight* award.
- An *insight* will be considered as piece of “knowledge” gained through the analysis of the hackathon data set. This could include, for example:
    - an interesting conclusion or understanding about the data
    - a recommendation or suggestion about how something could be changed or done differently, supported by data analysis
    - a link or association between different pieces of information in the data, particularly ones that are unexpected or non-intuitive
- Insights will be judged based upon how much practical impact the insight could have on the subject of the hackathon data set, and how well the associated analysis supports it.

Best Visualization

- Any team that presents a visualization as part of their final submission is eligible for the *Best Visualization* award.
- Visualizations can be either static graphics (e.g. a plot or 2-D graphic), or an interactive/dynamic visualization (e.g. a web page or a Shiny application with graphics components).
- Visualizations will be judged primarily on their ability to clearly convey an interesting and/or insightful aspect of the hackathon data set. Factors include:
    - how data visualizations are used to communicate interesting aspects of the data
    - the ability for someone not familiar with the data to understand (with the accompanying presentation) the main points of the visualization
    - the effective use of data graphics principles including, for example, the use of statistical summarizations, color, data subsetting, data highlighting, etc.
    - For static graphics, overly complex visualizations that detract from understanding the points of the visualization should be avoided.
    - for interactive/dynamic graphics, the interactive aspect should be a fundamental component to understanding the points of the visualization. Interactivity for interactivity-sake should be avoided.
    - interactive graphics will not be judged as “better” than static graphics simply because they are interactive.

Best Analysis

- Any team that presents results of their work that were obtained using statistical modeling or analysis will be eligible for the *Best Analysis* award.
- Any data modeling method can be used, including:
    - Linear regression, logistic regression, and other common statistical models
    - Machine learning models and algorithms such as decisions trees, Random Forest, support vector machines (SVM), and neural networks
- Teams presenting modeling results should also report an assessment of model performance (e.g. RMSE, confusion matrices/TPR/FPR/accuracy, ROC AUC, etc).

- Teams should also consider the use of “model explainers”, such as the R DALEX package, to visually show what’s happening with the models.
- Models will be judged based the underlying question the model is trying to address, the interpretation of the results, the performance of the model, and the validity of the methods used to derive the model.
- Alternatively the teams can perform some form of statistical analysis. This could be a descriptive statistics, summary statistics, an examination of the interrelations between features.
- This award is differentiated from the Best Insight award in that it is based on an overall assessment of the analysis. The Best Insight award is based on a single insight.

Best Presentation

- All teams that present their work will be eligible for the *Best Presentation* award.
- Presentation slides style.
    - Not too much information on each slide.
    - Visualization where it adds to the clarity.
- Presentation content.
    - Self-introduction.
    - Overview of presentation.
    - A clear story.
- Presentation clarity.
    - Short, sweet and to the point.
    - Easy to understand.
    - Covers the work completely.
    - Why? what? how?
    - No broad, sweeping statements.
- Presenter.
    - Clarity of speech.
    - Presentation style (engaging/boring).