

Student: Alan Danque

Class: DSC 640

Professor: Anthony Armstrong

Task 3: Blog Post

Git repository: <https://github.com/adanque/DSC640/tree/master/Task3-BlogPost>

Blogger URL: <https://fictiousairways.blogspot.com/>

For my task 3 blog, I designed my blog for a general audience. For this blog I wanted to combine the 3 phases of storytelling from Chapter 7 of Storytelling with Data with a high-level glimpse of the processes of CRISP-DM using the problem that Fictious Airways is faced with. Using the three phases of storytelling helped setup my blog nicely. After selecting a simplistic print friendly blog layout, picking out colors for my visuals per shades of blue for trustworthy good metrics, purples and pink for transitions, dark reds for bad, and selecting photos that support the transitional phases of my story, I started out with an inquisitively fun and catchy title. Then I started on the first phase of storytelling to setup the stage. In this first phase, I presented the environment – air travel what it is, who uses it, where is it used (supported by a map listing travel locations), when is it used and the protagonist of the story – Fictious Air. This was done using a likeable tone to relate to the needs and desires of my audience i.e. entrepreneurial travel, work related travel and the dream vacation travel. For the second phase, I introduced the team of data scientist that belongs to the protagonist. And then started building out the problem that our protagonist is going to be faced with which is air travel safety concerns causing leading to risk of bankruptcy. For the third phase, I generically described a few of the steps of CRISP-DM that our hero team of data scientists would use to create actionable steps to help remedy the problem at hand. That is - to understand the problem, understand the data, prepare the data, model, evaluate and deployed - which was inferred and not directly mentioned since this would be too technical for a general audience. What was used to infer these steps, was visualizations that represented survival counts vs fatality counts by aircraft so to display the comparison of how many fatalities and survivors there are per plane. A donut chart displaying the total count of crashes in the center of the donut. And a legend matching the arc colors for each aircraft listing their names and crash counts. This allowed for the crash count information by plane to be visibly relate-able with the following visuals. The next visual consisted of funnel chart that visualized the age of each aircraft. What was helpful about this visualization was the dark red bars to mark which planes had a survival rate less than 30%. Versus the age of the other planes that had a much higher survival rate in blue. This helped pave the way to the variables that illuminated the importance of the aggregated correlation analysis that was presented in a table that color code highlighted the mean survival rate for each aircraft. Again, using the dark red color to indicate which aircraft had a bad survival rate value. To further drive emphasis on the value of a metric called, "survival rate". A meter gauge visualization was presented to indicate the Fictious Airways overall survival rate. This helped create the idea of a measurable safety score for the company. And was something that can be used to identify and measure improvement after efforts were attempted to fix the conditions that may directly affect the problem. This was the goal for our data science protagonists' team. The data science team then leveraged this metric to present and prove that by decommissioning the three aircraft with less than a 30% survival rate would directly help fix the flight safety concern issue.