CS Rubric – Predicting Chicago Flight Delays

DS 4002 – Fall 2023
Due: December 15th, 2023
Submission Format: Upload a Github repository link with work to Canvas.

Individual Assignment

General Description: Submit a Github repository link with project work to Canvas.

Preparatory Assignments – None.

Why am I doing this? This opportunity will allow you to take what you have learned from previous data science courses and apply it to a meaningful project using real world data to create insights and contribute to necessary solutions. You will produce a final deliverable in the form of a presentation to a key audience using the insights you have gained from the coding and analysis you have completed, which will all be stored in a Github repository and submitted.

- Course Learning Objective: Clean and prepare given data for modeling and analysis
- Course Learning Objective: Utilize resources and prior knowledge to build a predictive model
- Course Learning Objective: Create an effective presentation for relevant audience from results

What am I going to do? You will first read the prompt and deliverable details included in the case study file. This will provide you with context behind the topic and data as well as a brief explanation of what is expected of you. You will then read the rest of this rubric to understand the specifics. You will then familiarize yourself with the data and create an analysis plan. The deliverables you will create include:

- Technical deliverables include data, relevant figures, source code, README file, and license file
- Executive deliverable is a presentation containing analytical insights for management team

All of this will be submitted electronically via a link to a Github repository built for this project.

Tips for success:

- Have fun with it. Don't restrict yourself to following a strict format. Think outside the box and see what you can do differently with the data.
- Ask for help. Don't hesitate to ask your peers, professor, or TA for help if you are stuck.
- Set time aside. Don't leave all of this work until the last minute. Make sure to set time aside so that you can produce your best work without the stress of having to rush.

How will I know I have Succeeded? You will meet expectations for this assignment when you follow the criteria in the rubric below.

Spec Category	Spec Details
Formatting	● Repository – A github repository that contains all project materials ○ Submit a link to the repository ○ Everything is contained in the repo or linked to it ○ Contents ■ A DATA folder ■ A FIGURES folder ■ A SRC folder ■ A PRESENTATION file ■ A README.md file ■ A LICENSE.md file ○ Upload presentation in PDF format ○ For code and data use appropriate format
DATA folder	● Goal: This folder contains all of the relevant data for this project ● Upload the data provided in the case study
FIGURES folder	● Goal: This folder contains all of the figures for this project ● Include any figures created during EDA, from the model results, or for the presentation in this folder ● Include relevant title and labels for each figure
SRC folder	● Goal: This folder contains all the source code for your project. ● Include all code files you produce during data cleaning, EDA, and model building
PRESENTATION file	● Goal: This file contains the presentation for the management team in PDF format ● It should use as little technical jargon possible (should be understandable to someone with little to no technical background) ● This presentation consists of: ○ Overview of Data ○ Key questions to answer ○ Model Overview (Brief explanation of how the model works – can use technical jargon in this part) ○ Results/Insights ○ Actionable recommendations to team
README.md file	● Goal: Introduce audience to case study content, code, and results ● This is a general overview of each entity in the Github repository ● Include a header for each entity listed in “Contents” in the Formatting category
LICENSE.md file	● Goal: This file explains the terms under which the repository may be used and cited. ● Select the MIT license from GitHub options on repository creation.