**Team Name:** Final Project Group 30

**Team Members:** 

Cole Anderton Jonathan Louis Cole Emerine

## **Project Title:**

Route Risk

**Problem:** What problem are we trying to solve? [0.25 point]

Have you ever wanted to see the route and potential traffic due to accidents before starting on a drive? Our application calculates the likelihood of being impacted by a car accident along a path between two cities?

**Motivation:** Why is this a problem? [0.25 point]

Many Americans drive throughout the country every single day. With a robust dataset on car accident reports, our user application aims to inform the user on the potential danger of a specific route they are interested in or for a journey they are about to embark on.

**Features**: When do we know that we have solved the problem? [0.25 point]

We will know that we have solved the problem when we can reliably get a probability statistic for a test set of city pairs.

**Data**: (Public data set we will be using and the link to the public data set) **or** (Schema of randomly generated data - i.e. what are the different columns in our dataset and the respective datatypes) [0.25 point]

Dataset includes data on time, duration, severity, location, and size (distance) of accidents. <u>US Accidents (2016 - 2023) (kaggle.com)</u>

**Tools:** Programming languages or any tools/frameworks we will be using [0.25 point] Python is the language that we decided as a group will work best. We aren't sure exactly which package we may utilize, but one tool we will potentially use is some form of geopy library that helps us to switch between city names and latitude and longitude values.

Visuals:/Sketches of the interface or the menu-driven program [0.25 points]



https://wireframe.cc/G6eKrS

## **Strategy**: Preliminary <u>algorithms</u> or <u>data structures</u> you may want to implement and how would you represent the data [0.25 points]

Take user input of a data structure from a dropdown and a pair of cities, and return a probability that the user will be impacted by an accident. Find a path from one city to the other using highway information and data processed from the dataset. Calculate probability of an accident occurring somewhere along the path. May additionally implement a start time so that probabilities could be made more accurate based on the time an accident is likely to occur in a city (if any) and the time and duration that the user would be in that city. A visualization of the data would be generated based on the cities along the path, which would contain data points about the accidents in each city. Our plan currently is to utilize a red-black tree and a b+ tree in our comparison. We will compare the efficiency/time complexity of these two trees within our application to see which is faster at reliably providing information for between two cities. The data will be represented through some form of visualization that shows the path between the two cities and the potential car accidents along the way.

## **Distribution of Responsibility and Roles:** Who is responsible for what? [0.25 points]

Role	Name
Visualization (of Dataset)	Jonathan Louis (PyOpenGL)
Data Structure Class	Cole Anderton, Cole Emerine, Jonathan Louis
User Input Processing	Cole Emerine
Probability Calculation	Cole Anderton
User Interface	Cole Emerine (PyGame)

## References

https://www.kaggle.com https://wireframe.cc/

https://pyopengl.sourceforge.net/